THE ELEMENTARY DIVISOR THEOREM FOR CERTAIN RINGS WITHOUT CHAIN CONDITION

OLAF HELMER

1. **Introduction.** The Elementary Divisor Theorem is known to hold in principal ideal rings and in rings with Euclidean algorithm.[1] It is an open question whether it holds in every *Prüfer ring*, that is, in every domain of integrity in which every ideal with finite basis is a principal ideal.[2] (A Prüfer ring can also be characterized as a domain of integrity in which the greatest common factor of any finite number of elements can be represented as a linear combination of these elements.) The Elementary Divisor Theorem will here be proved for what will be called "adequate" rings. They are special Prüfer rings, but not restricted by any equivalent to a chain condition, so that they comprise considerably more than just the principal ideal rings.[3]

2. **Definition of adequate rings.** Let R be a domain of integrity, a, b in R, and $a \neq 0$. By a *relatively prime part* of a with respect to b, written $RP(a, b)$, we shall understand a factor a_1 of a such that, if $a = a_1 \cdot a_2$,

(i) $(a_1, b) = 1$,

(ii) $(a_2, b) \neq 1$ for any non-unit factor a_3 of a_2.[4]

$RP(a, b)$ may or may not exist; if it does, it is, in a sense, a largest factor of a that is relatively prime to b.

We now define R to be an *adequate ring* if

(i) R is a Prüfer ring,

(ii) $RP(a, b)$ exists for all a, b in R with $a \neq 0$.

3. **Relationship to Prüfer rings and principal ideal rings.** By definition every adequate ring is a Prüfer ring. On the other hand, every

Received by the editors, May 29, 1942.

[1] B. L. van der Waerden, *Moderne Algebra*, vol. 2, Berlin, 1931, p. 122. For further reference, see the papers by J. H. M. Wedderburn (J. Reine Angew. Math. vol. 167 (1932)), N. Jacobson (Ann. of Math. (2) vol. 38 (1937)), and O. Teichmüller (Preuss. Akad. Wiss. Sitzungsber. 1937).

[2] Rings of this kind were considered by H. Prüfer in *Untersuchungen über Teilbarkeitseigenschaften in Körpern*, J. Reine Angew. Math. vol. 168 (1932).

[3] My thanks are due to Professor Reinhold Baer for several helpful suggestions.

[4] Notation throughout this paper: "$(a, b, \cdots)$" for "greatest common factor of $a, b, \cdots$," "$a \mid b$" for "a is a factor of b," "$a = b$" for "a equals b except possibly for a unit factor." (This last convention serves the purpose of replacing statements about principal ideals simply by statements about their generating elements.)

principal ideal ring is adequate; for in a ring of that kind we have unique prime power decomposition, and hence the existence of $RP(a, b)$ for $a \neq 0$ is obvious. An example of an adequate ring which is not a principal ideal ring is furnished by the set of integral functions with coefficients in a field F. That a ring of this type is a Prüfer ring has been proved in a previous paper.[5] The existence of $RP(a, b)$ is an immediate consequence of the unique factorization theorem for integral functions (Theorem 6 of the same paper). Theorem 8 of that paper, finally, states that the ring is not a principal ideal ring.

4. **Six lemmas.** Now let R be an adequate ring. Among the properties of such a ring are those stated in the six lemmas below. The first four are well known, and hold in any Prüfer ring; sketches of their proofs have been included, however, for reasons of completeness.

LEMMA 1. *If* $f \mid ab$ *and* $(f, a) = 1$, *then* $f \mid b$.

PROOF. Let $hf + ka = 1$; then $f \mid abk = b(1 - hf) = b - bhf$ and hence $f \mid b$.

LEMMA 2. *An element* a *has an inverse* a^* mod b,

$$a \cdot a^* \equiv 1 \pmod{b}, \tag{1}$$

if and only if $(a, b) = 1$. *Moreover,* a^*, *if it exists, is uniquely determined* mod b.

PROOF. If $aa^* \equiv 1 \pmod{b}$, then $aa^* = 1 + bk$ and hence $(a, b) = 1$. Conversely, if $(a, b) = 1$, then h and k exist such that $ha + kb = 1$, and hence $ah \equiv 1 \pmod{b}$ and $a^* = h$. Finally, if $aa^* \equiv aa^{**} \equiv 1 \pmod{b}$, then $a^* \equiv a^*(aa^{**}) \equiv (aa^*)a^{**} \equiv a^{**} \pmod{b}$.

LEMMA 3. *The congruences*

$$x \equiv a_i \pmod{m_i}, \qquad i = 1, 2, \cdots, n, \tag{2}$$

have a simultaneous solution x, *provided any two* m_i *are relatively prime. Moreover, the solution is uniquely determined* mod m, *where* $m = m_1 m_2 \cdots m_n$.

PROOF. Let $m = m_i M_i$. Since any two m_i are relatively prime it follows that $(m_i, M_i) = 1$. Hence, by Lemma 2, the M_i possess inverses M_i^* mod m_i. It is seen at once that

$$x \equiv a_1 M_1 M_1^* + \cdots + a_n M_n M_n^* \pmod{m} \tag{3}$$

[5] Theorem 9 in *Divisibility properties of integral functions*, Duke Math. J. vol. 6 (1940) p. 351.

is a required solution, and the uniqueness is obvious.

LEMMA 4. *Let $a_1, a_2, \cdots, a_n$ not be all zero, and*

$$(a_1, a_2, \cdots, a_n) = d; \tag{4}$$

then there exist elements a_{ij} $(i=2, \cdots, n; j=1, \cdots, n)$ such that

$$\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = d. \tag{5}$$

PROOF. If all but one a_k are zero, then $d=a_k$. Choose $a_{2k}=\cdots=a_{nk}=0$, and the unity matrix for the rest of the a_{ij}; the resulting determinant will be d or $-d$, if $-d$, a change of one 1 into -1 will put things right. If, on the other hand, no more than $n-2$ of the a_k are zero, we can proceed by induction. For $n=1$ the assertion is trivially correct. For $n>1$ let

$$d' = (a_2, \cdots, a_n); \tag{6}$$

by induction we can find a_{ij} $(i=2, \cdots, n-1; j=2, \cdots, n)$ such that

$$\begin{vmatrix} a_2 & \cdots & a_n \\ a_{22} & \cdots & a_{2n} \\ \cdot & \cdots & \cdot \\ a_{n-1,2} & \cdots & a_{n-1,n} \end{vmatrix} = d'. \tag{7}$$

Suppose now that

$$d = (a_1, d') = h \cdot a_1 + k \cdot d'. \tag{8}$$

We can then complete our determinant as follows:

$$\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ 0 & a_{n-1,2} & \cdots & a_{n-1,n} \\ (-1)^{n-1}k & (-1)^n h a_2/d' & \cdots & (-1)^n h a_n/d' \end{vmatrix} = h \cdot a_1 + k \cdot d' = d. \tag{9}$$

LEMMA 5. *$RP(a, b)$ is uniquely determined except for a unit factor.*

PROOF. Let

$$a = a_1 \cdot a_2 = a' \cdot a'', \tag{10}$$

where a_1 and a' are both largest factors of a that are relatively prime to b. We show first that $(a_2, a') = 1$. Suppose that $(a_2, a') = d \neq 1$; then $d \mid a_2$, and by definition of a_1, $(d, b) \neq 1$; but $d \mid a'$ implies $(d, b) \mid (a', b)$, so that $(a', b) \neq 1$, which is impossible by definition of a'. Hence $(a_2, a') = 1$, and we can now apply Lemma 1: since $a_2 \mid a'a''$, it follows that $a_2 \mid a''$. Similarly $a'' \mid a_2$, so that a_2 and a'' can only differ by a unit factor.

LEMMA 6. *Let* $RP(a, b) = r$, *and let* $d \mid a$ *with* $(d, r) = 1$; *then* $RP(d, b) = 1$.

PROOF. Let $a = r \cdot s$. By assumption and Lemma 1, $d \mid s$. Hence, by definition of r and s, $(\delta, b) \neq 1$ for every non-unit factor δ of d, and consequently $RP(d, b) = 1$.

5. **A preparatory theorem.** The main result of this paper will be a consequence of the following theorem.

THEOREM 1. *Let*

$$M = \begin{vmatrix} a_{11} & a_{12} \cdots a_{1k} \\ a_{21} & a_{22} \cdots a_{2k} \\ \cdot & \cdot \cdots \cdot \\ a_{n1} & a_{n2} \cdots a_{nk} \end{vmatrix} \tag{11}$$

be a matrix with coefficients in an adequate ring R, let

$$1 < n = \text{rank } M \leqq k \tag{12}$$

and

$$(a_{11}, a_{12}, \cdots, a_{1k}, a_{21}, \cdots, a_{nk}) = d; \tag{13}$$

then there exist $t_1, t_2, \cdots, t_{n-1}$ *in R such that*

$$(A_1, A_2, \cdots, A_k) = d, \tag{14}$$

where

$$A_i = a_{1i}t_1 + a_{2i}t_2 + \cdots + a_{n-1,i}t_{n-1} + a_{ni}, \quad i = 1, 2, \cdots, k. \tag{15}$$

PROOF. (a) Suppose the theorem has been proved for $d = 1$. Then the general case results as follows. Let

$$N = (1/d) \cdot M; \tag{16}$$

the coefficients

$$b_{ij} = a_{ij}/d \tag{17}$$

of the matrix N are then relatively prime, and the theorem is considered proved for this case; there are, therefore, quantities $t_1, t_2, \cdots, t_{n-1}$ in R such that

$$(B_1, B_2, \cdots, B_k) = 1, \tag{18}$$

where

$$B_i = b_{1i}t_1 + b_{2i}t_2 + \cdots + b_{n-1,i}t_{n-1} + b_{ni} = A_i/d. \tag{19}$$

Hence

$$(A_1, A_2, \cdots, A_k) = d\cdot(B_1, B_2, \cdots, B_k) = d. \tag{20}$$

(b) For $d=1$ the proof will be given by induction with respect to n. First, let $n=2$. We have to find t so that the

$$A_i = a_{1i}t + a_{2i}, \qquad i = 1, 2, \cdots, k, \tag{21}$$

have the greatest common factor 1. Note that any common factor of A_i and A_j is also a factor of

$$D_{ij} = \begin{vmatrix} a_{1i} & a_{1j} \\ a_{2i} & a_{2j} \end{vmatrix} = a_{1i}A_j - a_{1j}A_i. \tag{22}$$

If, therefore, we set

$$D = (D_{12}, D_{13}, D_{23}, D_{14}, \cdots, D_{k-1,k}), \tag{23}$$

then, however t be chosen,

$$(A_1, A_2, \cdots, A_k) \mid D. \tag{24}$$

On account of (12), not all D_{ij} are zero; hence $D\neq 0$, and we may define:

$$\begin{cases} D_1 = RP(D, a_{11}), \\ D_2 = RP\left(\dfrac{D}{D_1}, a_{12}\right), \\ \cdots\cdots\cdots\cdots, \\ D_k = RP\left(\dfrac{D}{D_1D_2\cdots D_{k-1}}, a_{1k}\right). \end{cases} \tag{25}$$

We then have, obviously,

$$(D_i, D_j) = 1, \qquad i \neq j, \tag{26}$$

because a largest relatively prime factor of a quantity is always prime to the supplementary factor. Consequently, by Lemma 3, there exists a simultaneous solution t of the congruences

$$t \equiv a_{1i}^{*}\cdot(1 - a_{2i}) \pmod{D_i}, \qquad i = 1, 2, \cdots, k, \tag{27}$$

where a_{1i}^{*} is an inverse of a_{1i} mod D_i:

$$a_{1i}^{*}\cdot a_{1i} \equiv 1 \pmod{D_i} \tag{28}$$

(the existence of a_{1i}^{*} being guaranteed by Lemma 2, since $(a_{1i}, D_i) = 1$ by virtue of (25)). With t chosen thus, we shall now show that

$$(A_1, A_2, \cdots, A_k) = 1. \tag{29}$$

From (21) and (27) we have

$$A_i = a_{1i}t + a_{2i} \equiv 1 \pmod{D_i}. \tag{30}$$

Suppose, now, that the A_i have some non-unit factor c in common; then, by (24), it would follow that $c \mid D$. This, together with (30), implies that

$$(c, D_i) = 1; \tag{31}$$

this means that c is a factor of D that is prime to all D_i, and we can apply Lemma 6, with the result that

$$RP(c, a_{1i}) = 1, \qquad i = 1, 2, \cdots, k. \tag{32}$$

But then, as c was not supposed to be a unit, none of its non-unit factors is prime to any a_{1i}. In particular,

$$\begin{cases} (c, a_{11}) = c_1 \neq 1, \\ (c_1, a_{12}) = c_2 \neq 1, \\ \cdots\cdots\cdots, \\ (c_{k-1}, a_{1k}) = c_k \neq 1; \end{cases} \tag{33}$$

the a_{1i} thus have a common non-unit factor c_k, and, since $c_k \mid A_i$, it follows that c_k is also a factor of $A_i - a_{1i}t = a_{2i}$. But this is incompatible with the assumption that $d = 1$.

(c) Suppose now that the theorem has been proved for values less than n; we then have to show that it holds for n provided $d = 1$. Let

$$d_1 = (a_{21}, a_{22}, \cdots, a_{2k}, a_{31}, \cdots, a_{nk}); \tag{34}$$

applying induction, we can find $t_2, t_3, \cdots, t_{n-1}$ so that

$$(B_1, B_2, \cdots, B_k) = d_1 \tag{35}$$

where

$$B_i = a_{2i}t_2 + a_{3i}t_3 + \cdots + a_{n-1,i}t_{n-1} + a_{ni}, \qquad i = 1, 2, \cdots, k, \tag{36}$$

and it remains to determine t_1 suitably. We have

$$A_i = a_{1i}t_1 + B_i \tag{37}$$

with

$$\begin{aligned}(a_{11}, B_1, a_{12}, B_2, \cdots, a_{1k}, B_k) &= (a_{11}, a_{12}, \cdots, a_{1k}, d_1) \\ &= (a_{11}, a_{12}, \cdots, a_{1k}, a_{21}, \cdots, a_{nk}) = 1.\end{aligned} \tag{38}$$

Therefore, using induction with $n=2$, we can find t_1 such that

$$(A_1, A_2, \cdots, A_k) = 1, \tag{39}$$

and the proof of Theorem 1 is completed.

6. **The case of a square matrix with nonzero determinant.** Before taking up the general case of the Elementary Divisor Theorem we shall first treat the case of a nonsingular matrix.

THEOREM 2. *Let R be an adequate ring, M a square matrix with coefficients in R, and $|M| \neq 0$. Then there exist two matrices X and Y with coefficients in R, such that*

$$|X| = |Y| = 1 \tag{40}$$

and

$$XMY = \begin{pmatrix} e_1 & 0 & \cdots & 0 \\ 0 & e_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & e_n \end{pmatrix}, \tag{41}$$

where the e_i are the elementary divisors of M.

PROOF. For $n=1$ the theorem is trivially true. For $n>1$ assume the theorem proved for all values less than n. It is sufficient to consider the case where e_1 (which is the greatest common factor of the coefficients of M) equals 1; for otherwise we first deal with the matrix M/e_1, whose elementary divisors are $1, e_2/e_1, \cdots, e_n/e_1$; hence if

$$X\frac{M}{e_1}Y = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & e_2/e_1 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & e_n/e_1 \end{pmatrix}, \tag{42}$$

we have at once

$$XMY = \begin{pmatrix} e_1 & 0 & \cdots & 0 \\ 0 & e_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & e_n \end{pmatrix}. \tag{43}$$

Let now $e_1=1$ and $M=(a_{ij})$. Applying Theorem 1, with $k=n$ and $d=1$, we can find $t_1, t_2, \cdots, t_{n-1}$ in R such that

$$(A_1, A_2, \cdots, A_n) = 1, \tag{44}$$

where

$$A_i = a_{1i}t_1 + a_{2i}t_2 + \cdots + a_{n-1,i}t_{n-1} + a_{ni}, \qquad i = 1, 2, \cdots, n. \tag{45}$$

Applying Lemma 4 to these A_i, a square matrix A with coefficients in R can be found whose first row is $A_1, A_2, \cdots, A_n$ and whose determinant is 1. If h_i is the minor belonging to A_i, we have

$$h_1A_1 + h_2A_2 + \cdots + h_nA_n = 1. \tag{46}$$

As $|A|=1$, we can form, in R, the inverse matrix

$$Q = A^{-1}. \tag{47}$$

Note that the first column of Q is $h_1, h_2, \cdots, h_n$. If, now, we set

$$q_i = h_1a_{i1} + h_2a_{i2} + \cdots + h_na_{in}, \qquad i = 1, 2, \cdots, n, \tag{48}$$

we can define

$$P = P_1 \cdot P_2, \tag{49}$$

where

$$P_1 = \begin{pmatrix} 1 & 0 \cdots & 0 \\ q_1 & 0 \cdots & -1 \\ \cdot & \cdots & \cdot \\ q_{n-1} & -1 \cdots & 0 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} t_1 \cdots t_{n-1} & 1 \\ 0 \cdots 1 & 0 \\ \cdot \cdots \cdot & \cdot \\ 1 \cdots 0 & 0 \end{pmatrix}. \tag{50}$$

Then

$$|P_1| = |P_2| = (-1)^{n(n-1)/2} \tag{51}$$

so that

$$|P| = |P_1| \cdot |P_2| = 1. \tag{52}$$

Multiplying out, we obtain

$$P = \begin{pmatrix} t_1 & t_2 & \cdots & t_{n-1} & 1 \\ q_1t_1 - 1 & q_1t_2 & \cdots & q_1t_{n-1} & q_1 \\ q_2t_1 & q_2t_2 - 1 & \cdots & q_2t_{n-1} & q_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ q_{n-1}t_1 & q_{n-1}t_2 & \cdots & q_{n-1}t_{n-1} - 1 & q_{n-1} \end{pmatrix}. \tag{53}$$

Notice that we also have

$$|Q| = 1. \tag{54}$$

We now multiply M from the left by P and from the right by Q:

$$PMQ = \begin{pmatrix} A_1 & A_2 & \cdots & A_n \\ q_1A_1 - a_{11} & q_1A_2 - a_{12} & \cdots & q_1A_n - a_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ q_{n-1}A_1 - a_{n-1,1} & q_{n-1}A_2 - a_{n-1,2} & \cdots & q_{n-1}A_n - a_{n-1,n} \end{pmatrix}$$

$$\cdot \begin{pmatrix} h_1 & * & \cdots & * \\ h_2 & * & \cdots & * \\ \cdot & \cdot & \cdots & \cdot \\ h_n & * & \cdots & * \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \cdot & \cdot & \cdots & \cdot \\ 0 & * & \cdots & * \end{pmatrix} \tag{55}$$

where the zeros in the first row of this matrix arise from multiplying the row $A_1, A_2, \cdots, A_n$ by a row of minors of A not belonging to that row, while the zero in the $(i+1)$st place of the first column results as follows:

$$\begin{aligned} &(q_iA_1 - a_{i1})h_1 + \cdots + (q_iA_n - a_{in})h_n \\ &\quad = q_i(A_1h_1 + \cdots + A_nh_n) - (a_{i1}h_1 + \cdots + a_{in}h_n) \\ &\quad = q_i \cdot 1 - q_i = 0. \end{aligned} \tag{56}$$

We can now apply induction to the matrix M_1 of order $n-1$ that appears in the lower right corner of PMQ. Let

$$X_1M_1Y_1 = \begin{pmatrix} e_1' & 0 & \cdots & 0 \\ 0 & e_2' & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & e_{n-1}' \end{pmatrix} \tag{57}$$

where $|X_1| = |Y_1| = 1$ and where $e_1', e_2', \cdots, e_{n-1}'$ are the elementary divisors of M_1. Then, if we set

$$(58)\qquad X = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & X_1 \\ 0 & \end{pmatrix} \cdot P, \qquad Y = Q \cdot \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Y_1 \\ 0 & \end{pmatrix},$$

we obtain immediately

$$(59)\qquad XMY = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e_1' & 0 & \cdots & 0 \\ 0 & 0 & e_2' & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \\ 0 & 0 & 0 & \cdots & e_{r-1}' \end{pmatrix},$$

and all that remains to be shown is that

$$(60)\qquad e_i = e_{i-1}', \qquad i = 2, 3, \cdots, n.$$

Now, clearly, the determinant divisors of XMY are $1, e_1', e_1'e_2', \cdots, e_1'e_2' \cdots e_{n-1}'$; hence its elementary divisors are $1, e_1', e_2', \cdots, e_{n-1}'$, and the rest follows from the fact that "equivalent" matrices (here M and XMY) have the same elementary divisors (proof literally as for the ring of rational integers).

7. **The general case.** We can now prove the general Elementary Divisor Theorem:

THEOREM 3. *Let R be an adequate ring, and M a matrix of type (m, n) with coefficients in R. Then there exist two matrices X and Y with coefficients in R and of types (m, m) and (n, n) respectively, such that*

$$(61)\qquad |X| = |Y| = 1$$

and

$$(62)\qquad XMY = \begin{pmatrix} e_1 & 0 & & \cdots & 0 \\ 0 & e_2 & \ddots & & \cdot \\ \cdot & \ddots & \ddots & & \cdot \\ & & & e_r & \\ \cdot & & & & \cdot \\ & & & 0 & \\ \cdot & & & & \ddots \\ 0 & \cdot & & \cdots & 0 \end{pmatrix}$$

where $e_1, e_2, \cdots, e_r$ are the elementary divisors of M.

PROOF. (a) Let $M=(a_{ij})$, of rank r. If $r=0$, we can set $X=E_m$ and $Y=E_n$. Otherwise not all a_{ij} are zero, and we have

$$(a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, \cdots, a_{mn}) = d \neq 0. \tag{63}$$

The case where $m=n=r$ is covered by Theorem 2. If $r<m$ or $r<n$ or both, it will be shown below that M can be transformed unitarily into a matrix of the form

$$\begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix} \tag{64}$$

where N is of type (r, r) (and hence $|N| \neq 0$). Once this has been done, Theorem 2 can be applied to N; and if, say, X_1 and Y_1 take N into the desired normal form, then

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & E_{m-r} \end{pmatrix}, \quad Y = \begin{pmatrix} X_2 & 0 \\ 0 & E_{n-r} \end{pmatrix} \tag{65}$$

will do the same for M.

(b) Let $r<m$. The m rows of M must then be linearly dependent in R:

$$j_1 a_{1i} + j_2 a_{2i} + \cdots + j_m a_{mi} = 0, \qquad i = 1, 2, \cdots, n, \tag{66}$$

where the j_i may be assumed to be relatively prime. By Lemma 4 we can find elements j_{ik} in R $(i=1, \cdots, m-1; k=1, \cdots, m)$ such that the matrix

$$J = \begin{pmatrix} j_{11} & j_{12} \cdots & j_{1m} \\ \cdot & \cdot \cdots & \cdot \\ j_{m-1,1} & \cdot \cdots & j_{m-1,m} \\ j_1 & j_2 \cdots & j_m \end{pmatrix} \tag{67}$$

has determinant 1; and if we multiply M on the left by J, we obtain a matrix whose last row consists of zeros.

Clearly, if $r<n$, an analogous multiplication of M on the right by a unitary matrix will generate a last column of zeros.

This procedure can be applied repeatedly. If

$$X_1 M X_2 = \begin{pmatrix} M_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{68}$$

where M_1 is of type (m_1, n_1) with $r<m_1$ or $r<n_1$, a further row or column of M_1 (and hence of M) can be turned into zeros. This process will break off only when the form

$$\begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix} \tag{69}$$

with N of type (r, r) has been reached. This, according to (a), completes the proof.

COLLEGE OF THE CITY OF NEW YORK

THE GEOMETRY OF VELOCITY SYSTEMS

EDWARD KASNER AND JOHN DeCICCO

1. **Introduction.** In connection with the investigation of the differential geometric properties of positional fields of force, Kasner introduced certain important systems of curves which he termed *velocity systems*.[1] In this paper, we propose to present some of the old and many new geometric properties of any velocity system.

Since velocity systems serve to characterize the conformal group, many of our results are logically theorems of conformal geometry. This may be contrasted with the study of all the trajectories of positional fields of force, which is essentially differential projective geometry.

2. **Velocity systems.** We shall find it most convenient for our study to use the minimal coordinates $u=x+iy$, $v=x-iy$, $p=dv/du$, $q=d^2v/du^2$, where, of course, (x, y) are the cartesian coordinates of the plane.

In the plane, consider a particle of unit mass moving in any field of force whose components parallel to the u-axis and the v-axis are $d(u, v)$ and $c(u, v)$. The equations of motion are

$$\frac{d^2u}{dt^2} = d(u, v), \qquad \frac{d^2v}{dt^2} = c(u, v), \tag{1}$$

where t is the time. Now if r is the radius of curvature along a trajectory, we have

$$\omega^2 = rN, \tag{2}$$

Presented to the Society, February 22, 1941, under the title *Conformal geometry of velocity systems*; received by the editors June 23, 1942.

[1] Kasner, *Differential-geometric aspects of dynamics*, Amer. Math. Soc. Colloquium Publications vol. 3 (1912); Trans. Amer. Math. Soc. vols. 9–10 (1908–1909).

where $\omega = p^{1/2}du/dt$ is the velocity (speed) and N is the principal normal component of the force given by

$$N = \frac{c - dp}{2ip^{1/2}}. \tag{3}$$

Since (1) defines a three-parameter family of curves, the path is uniquely determined if the initial position, the initial direction, and the initial velocity are given.

If only the initial position and the initial direction of motion are given, then the initial radius of curvature will depend on the initial velocity. In the actual trajectory, the velocity ω varies from point to point. If now the square of the velocity ω^2 is replaced by some constant, say 1/const., the resulting equation is

$$q = \text{const. } p(c - dp). \tag{4}$$

The curves satisfying this differential equation, they are not in general trajectories, are called *velocity curves.*

For any field, a curve is a velocity curve corresponding to the speed ω_0, provided any particle starting from a lineal element of the curve with that speed, describes a trajectory which is initially osculated by that curve. In a given field of force, there are in all ∞^3 trajectories and ∞^3 velocity curves.

If the constant is given, there are ∞^2 velocity curves. In particular, if the constant and hence the velocity is taken to be unity, the following result is obtained.

THEOREM 1. *Any system of ∞^2 curves defined by a differential equation of the form*[2]

$$q = p(c - dp), \tag{5}$$

may be regarded as the totality of velocity curves corresponding to unit velocity in the unique field of force whose u-component is $d(u, v)$ and whose v-component is $c(u, v)$.

Throughout this paper, we shall understand a velocity system to be the totality of velocity curves with unit velocity. That is, any velocity system is given by the preceding differential equation.

THEOREM 2. PROPERTY A. *A set of ∞^2 curves of the plane is a velocity system if and only if the ∞^1 osculating circles of the ∞^1 curves of the set passing through any point (u, v), constructed at (u, v), form a linear pencil of circles.*

[2] The corresponding cartesian equation of the general velocity system is $y'' = (1+y'^2)(\chi - y'\phi)$ where ϕ and χ are any functions of (x, y).

Of course, the centers of these circles describe a straight line. The other point of intersection (U, V) of this pencil of circles is

$$U = u + 2/c, \qquad V = v + 2/d. \tag{6}$$

We shall call this the *associated correspondence* of the velocity system.

Obviously any arbitrary point transformation except identity, may serve as the associated correspondence of a unique velocity system.

Special important types of velocity systems are natural families, isogonal families, and Γ families of curves. Now we shall consider these sets of curves (special wexes).

3. **Natural families.** A natural family consists of the extremals connected with the variation problem of the form

$$\int F(u, v)ds = \text{minimum}, \tag{7}$$

where F is a point function and $ds = p^{1/2}du$ is the element of arc. Any natural family is a velocity system for which $c_v = d_u$.

Kasner has already indicated the many problems which are connected with natural families. Thus natural families arise in the study of (a) the trajectories of a conservative field of force, (b) general catenaries, (c) the paths of light in an isotropic medium, and (d) the geodesics of a surface.[1] (See papers by Lipka, Douglas, Blaschke, Fialkow, DeCicco, Struik.)

A natural family of curves is a special type of velocity system. Hence in addition to the Property A, it must possess an additional differential property.

THEOREM 3. PROPERTY B. *At any point (u, v), there exist two lineal elements E_1 and E_2, each of which is converted by the associated correspondence of the velocity system into a cocircular element at the point (U, V). Our velocity system is a natural family if and only if E_1 and E_2 are orthogonal.*

Thus any system of ∞^2 curves with the properties A and B is a natural family.

4. **Isogonal families.** Any nonminimal simple family of ∞^1 curves is given by a differential equation of the first order $p = e^{f(u,v)}$. The set of *all* simple families of curves each of which makes a constant angle with all the curves of the given simple family is called a complete *isogonal family*. Any isogonal family is a velocity system for which $c = f_u$ and $d = -f_v$ so that $c_v = -d_u$.

THEOREM 4. PROPERTY C. *At any point* (u, v), *there exist two lineal elements* E_1 *and* E_2, *each of which is converted by the associated correspondence of the velocity system into an element orthogonal to the cocircular element at the point* (U, V). *Our velocity system is an isogonal family if and only if* E_1 *and* E_2 *are othogonal.*

Thus the conditions necessary and sufficient for an isogonal family are properties A and C. Thus the famous theorem of Cesàro-Scheffers is completed.

5. Γ_0 **systems.** The search for those velocity systems which are both natural and isogonal yields the following result.

THEOREM 5. *For a velocity system to be both natural and isogonal (that is, to possess properties* A, B, *and* C), *it is necessary and sufficient that it be the complete set of isogonal trajectories of an isothermal family.*

This family may be called a *conformal rectilinear wex.*[3] That is, it is conformally equivalent to the ∞^2 straight lines of the plane. Any such system shall be denoted by Γ_0.[4] Any system Γ_0 may be characterized geometrically as a velocity system whose associated point correspondence is direct conformal.

6. **The** Γ **families.** Any set of ∞^2 curves which is conformally equivalent to the set of ∞^2 circles orthogonal to a fixed proper circle (or a fixed straight line) is called a Γ *family.*[5]

The finite form of a Γ family is

$$a_0\phi(u)\psi(v) + b_0[\phi(u) + \psi(v)] + c_0 = 0. \tag{8}$$

Upon eliminating the constants, a_0, b_0, c_0, we discover that a Γ family is a velocity system where the basic functions

$$c = \frac{\phi_{uu}}{\phi_u} - \frac{2\phi_u}{\phi - \psi}, \qquad d = \frac{\psi_{vv}}{\psi_v} + \frac{2\psi_v}{\phi - \psi}. \tag{9}$$

Now we proceed to find the conditions on the functions c and d that a velocity system be a Γ family. Solving the first equation for ψ,

[3] Kasner has introduced the word *wex* to represent any system of ∞^2 curves. Thus a wex is given by a differential equation of the second order $y'' = f(x, y, y')$. A web denotes $n\,\infty^1$ curves.

[4] The authors have determined all curvature element transformations which preserve the class of Γ_0 systems. See Kasner and DeCicco, *Transformation theory of the isogonal trajectories of an isothermal family*, Proc. Nat. Acad. Sci. U.S.A., vol. 28 (1942).

[5] The Γ and Γ_0 systems are special subsets of Ω families. Any Ω family is conformally equivalent to the ∞^3 circles. See Kasner and DeCicco, *Families of curves conformally equivalent to circles*, Trans. Amer. Math. Soc. vol. 49 (1941) pp. 378–391.

and the second for ϕ, we find

$$\psi = \phi + \frac{2\phi_u}{c - \phi_{uu}/\phi_u}, \qquad \phi = \psi + \frac{2\psi_v}{d - \psi_{vv}/\psi_v}. \tag{10}$$

By differentiating the first equation of (10) with respect to u and the second with respect to v, we obtain

$$2c_u - c^2 = \frac{2\phi_u\phi_{uuu} - 3\phi_{uu}^2}{\phi_u^2}, \quad 2d_v - d^2 = \frac{2\psi_v\psi_{vvv} - 3\psi_{vv}^2}{\psi_v^2}. \tag{11}$$

Substituting the values of ϕ and ψ as determined by (10) into these equations, we find

$$dc_v = c_{vv}, \qquad cd_u = d_{uu}. \tag{12}$$

THEOREM 6. PROPERTY D. *A velocity system represents a* Γ *family of curves if and only if the functions* c *and* d *with* $c_v \neq 0$ *and* $d_u \neq 0$ *satisfy either one of the three equivalent systems of partial differential equations of the second order.*

$$c_{uv} = cc_v, \qquad c_{vv} = dc_v; \tag{13.1}$$

$$d_{uv} = dd_u, \qquad d_{uu} = cd_u; \tag{13.2}$$

$$c_v = d_u, \qquad c_{uv} = cc_v, \qquad d_{uv} = dd_u. \tag{13.3}$$

The last set of equations demonstrates that every Γ family is a natural family. But the converse is not true. The only possible Γ families which are isogonal are the Γ_0 systems.

7. **Velocity systems which possess isothermal families.** We shall now classify velocity systems according to the number of isothermal families they contain. For the velocity system (5) to contain the simple family $p = e^{f(u,v)}$, it is necessary and sufficient that

$$f_u + e^f f_v = c - e^f d. \tag{14}$$

If the simple family is isothermal, then, as deduced from the theorem of Lie, $f_{uv} = 0$. Differentiating the above partially with respect to v and simplifying, we find

$$f_{vv} + f_v^2 + df_v = e^{-f}c_v - d_v. \tag{15}$$

Differentiating this partially with respect to u and making use of (14), we obtain

$$(c_v - d_u)f_v = (d_{uv} - dc_v) - (c_{uv} - cc_v)e^{-f}. \tag{16}$$

From these equations, it is concluded that a velocity system will contain *exactly* ∞^2 isothermal systems if and only if it is a Γ family. This is a characteristic property of Γ families (including Γ_0).

Suppose next our velocity system is natural but *not* a Γ family. Then there exists possibly *one* isothermal family of curves. This possible isothermal family is

$$p = \frac{c_{uv} - cc_v}{d_{uv} - dc_v} \quad \text{where} \quad c_v = d_u. \tag{17}$$

Finally let us suppose that our family is *not* natural. Such a velocity system may possess at most ∞^1 isothermal families. These are defined by (14) and (16).

Differentiating the equation (16) with respect to u and eliminating f_u and f_v from the result, we find

$$e' = \frac{\left(\frac{\partial}{\partial u}\,\frac{c_{uv} - cc_v}{c_v - d_u}\right) - \left(\frac{c_{uv} - cc_v}{c_v - d_u}\right)^2 - \frac{c(c_{uv} - cc_v)}{c_v - d_u}}{\left(\frac{\partial}{\partial u}\,\frac{d_{uv} - dc_v}{c_v - d_u}\right) - \frac{(c_{uv} - cc_v)(d_{uv} - dc_v)}{(c_v - d_u)^2} - \frac{d(c_{uv} - cc_v)}{c_v - d_u}}. \tag{18}$$

Again by differentiating (16) with respect to v and eliminating f_u, f_v, f_{vv} from the result, we find

$$e' = -\frac{\left(\frac{\partial}{\partial v}\,\frac{c_{uv} - cc_v}{c_v - d_u}\right) + \frac{(c_{uv} - cc_v)(d_{uv} - dc_v)}{(c_v - d_u)^2} + \frac{d(c_{uv} - cc_v)}{c_v - d_u} + c_v}{\left(\frac{\partial}{\partial v}\,\frac{d_{uv} - dc_v}{c_v - d_u}\right) + \left(\frac{d_{uv} - dc_v}{c_v - d_u}\right)^2 + \frac{d(d_{uv} - dc_v)}{c_v - d_u} + d_v}. \tag{19}$$

THEOREM 7. *A velocity system of curves may contain exactly* ∞^2, ∞^1, *one, or no isothermal systems as described below.*

(1) A velocity system will contain *exactly* ∞^2 isothermal families if and only if it is a Γ family (including Γ_0).

(2) A nonnatural velocity system will contain *exactly* ∞^1 isothermal systems if and only if the functions c and d satisfy the four partial differential equations of the third order obtained by setting equal to zero the numerators and denominators of the fractions in (18) and (19).

(3) A nonnatural velocity system will contain exactly *one* isothermal system if the functions c and d satisfy one partial differential equation of the third order (obtained by the equalities of the right-

hand sides of (18) and (19), one of the fourth order (the family (18) or (19) lies in the velocity system), and one of the fifth order (the family (18) or (19) is isothermal).

A natural family will contain exactly one isothermal family if the functions c and d satisfy identically one partial differential equation of third order (the family (17) must lie in the velocity system), and another of fourth order (the family (17) must be isothermal).

(4) All other types of velocity systems contain no isothermal families. This is of course the general case.

THEOREM 8. PROPERTY E. *A velocity system is a* Γ *or* Γ_0 *family if and only if it contains* ∞^2 *isothermal families.*

Properties A and E are necessary and sufficient conditions for the characterization of a Γ or Γ_0 family.

8. **Reciprocal velocity systems.** Let Σ be the associated point correspondence (6) of a velocity system S. The unique velocity system S^* whose associated point correspondence is the inverse Σ^{-1} of Σ is said to be the *reciprocal*[6] or *conjugate* of S.

By (7), we find that the reciprocal velocity system S^* of S is

$$Q = P(\gamma - \delta P), \tag{20}$$

where

$$\gamma = -c, \qquad \delta = -d,$$

$$\frac{\partial \gamma}{\partial U} = -\frac{c^2}{J}\left[c_u(d^2 - 2d_v) + 2c_v d_u\right], \qquad \frac{\partial \gamma}{\partial V} = -\frac{c^2 d^2 c_v}{J},$$

$$\frac{\partial \delta}{\partial U} = -\frac{c^2 d^2 d_u}{J}, \qquad \frac{\partial \delta}{\partial V} = -\frac{d^2}{J}\left[2d_u c_v + d_v(c^2 - 2c_u)\right], \tag{21}$$

$$J = (c^2 - 2c_u)(d^2 - 2d_v) - 4d_u c_v.$$

THEOREM 9. *The reciprocal velocity system* S^* *of a given velocity system* S *is natural or isogonal according as* S *is natural or isogonal.*

The reciprocity is not valid in general for Γ systems (of course excluding Γ_0). We now wish to find those particular velocity systems whose reciprocals are Γ families.

Before proceeding, let us note the following formulas in addition to those given by (21)

---

[6] Scheffers, Math. Ann. vol. 60 (1905) pp. 491–531.

$$\frac{\partial^2\gamma}{\partial U\partial V}\Big/\frac{\partial\gamma}{\partial V}=\frac{2c}{J}\left[c_u(d^2-2d_v)+2c_vd_u\right]+\frac{2c^2dd_u}{J}$$
$$+\frac{c^2}{Jc_v}\left[c_{uv}(d^2-2d_v)+2d_uc_{vv}\right]$$
$$-\frac{c^2}{J^2}\left[J_u(d^2-2d_v)+2J_vd_u\right],$$

(22)

$$\frac{\partial^2\delta}{\partial U\partial V}\Big/\frac{\partial\delta}{\partial U}=\frac{2cd^2c_v}{J}+\frac{2d}{J}\left[2c_vd_u+d_v(c^2-2c_u)\right]$$
$$+\frac{d^2}{Jd_u}\left[2c_vd_{uu}+d_{uv}(c^2-2c_u)\right]$$
$$-\frac{d^2}{J^2}\left[2J_uc_v+J_v(c^2-2c_u)\right].$$

By (21) and (22), we may prove the following result.

THEOREM 10. *A velocity system S is the reciprocal of a Γ family S^* if and only if its functions c and d with $c_v\neq 0$ and $d_u\neq 0$ satisfy the partial differential equations of third order*

$$c_v=d_u,\qquad \frac{J_u}{J}=c+\frac{c_{uv}}{c_v},\qquad \frac{J_v}{J}=d+\frac{c_{vv}}{c_v}. \tag{23}$$

In the *real* domain, it may be proved that the only velocity systems such that both S and S^* are Γ families must consist of the ∞^2 circles orthogonal to a fixed circle. In that event, S and S^* coincide.

9. **The conformal covariants of a velocity system.**[7] Upon extending a conformal transformation twice, we find the formulas

$$U=\phi(u),\ V=\psi(v),\ P=\frac{\psi_v}{\phi_u}p,\ \frac{Q}{P}=\frac{q}{\phi_up}+\frac{\psi_{vv}p}{\phi_u\psi_v}-\frac{\phi_{uu}}{\phi_u^2}, \tag{24}$$

which have many applications.

THEOREM 11. *Every velocity system $S(c, d)$ is converted into a velocity system $\Sigma(C, D)$ by a conformal transformation. The transformation formulas between corresponding velocity systems S and Σ are*

[7] For the analogous conformal geometry of a simple family of curves, see DeCicco, *The two conformal covariants of a field*, Revista di Matematica vol. 2 (1941) pp. 59–66. A paper on *Isothermal families* will appear in the forthcoming Pastor commemoration volume edited by Beppo Levi.

$$C = \frac{c}{\phi_u} - \frac{\phi_{uu}}{\phi_u^2}, \qquad D = \frac{d}{\psi_v} - \frac{\psi_{vv}}{\psi_v^2}. \tag{25}$$

(It may be proved that if a quasi-contact transformation of curvature elements carries every velocity system into a velocity system, then it must be a conformal transformation.)

We now inquire: what are the necessary and sufficient conditions that the velocity systems $S(c, d)$ and $\Sigma(C, D)$ be equivalent by a conformal transformation? That is, given the four functions $c(u, v)$, $d(u, v)$, $C(\phi, \psi)$, $D(\phi, \psi)$, can we find $\phi(u)$ and $\psi(v)$ such that the equations (25) are satisfied?

Firstly upon differentiating the first equation (25) with respect to v and the second with respect to u, we find

$$\phi_u \psi_v = \frac{c_v}{C_\psi} = \frac{d_u}{D_\phi}. \tag{26}$$

Differentiating this with respect to u and v and eliminating ϕ_{uu} and ψ_{vv} by (25), we obtain

$$\phi_u = \left(\frac{c_{uv}}{c_v} - c\right) \Big/ \left(\frac{C_{\phi\psi}}{C_\psi} - C\right), \quad \psi_v = \left(\frac{d_{uv}}{d_u} - d\right) \Big/ \left(\frac{D_{\phi\psi}}{D_\phi} - D\right). \tag{27}$$

Now these equations will determine appropriate functions ϕ_u and ψ_v if and only if the derivative of the first and second fractions with respect to v and u, respectively, are zero. Accordingly we find that the velocity system $S(c, d)$ possesses the two absolute covariants

$$\left[\frac{\partial}{\partial v} \log\left(\frac{c_{uv}}{c_v} - c\right)\right] \Big/ \left(\frac{d_{uv}}{d_u} - d\right) \quad \text{and}$$
$$\left[\frac{\partial}{\partial u} \log\left(\frac{d_{uv}}{d_u} - d\right)\right] \Big/ \left(\frac{c_{uv}}{c_v} - c\right). \tag{28}$$

But these values as given by (27) must also satisfy the equations (25) and (26). Hence our velocity system $S(c, d)$ possesses the three absolute covariants

$$\left(\frac{c_{uv}}{c_v} - c\right)\left(\frac{d_{uv}}{d_u} - d\right) \Big/ c_v,$$
$$\left[c - \frac{\partial}{\partial u} \log\left(\frac{c_{uv}}{c_v} - c\right)\right] \Big/ \left(\frac{c_{uv}}{c_v} - c\right) \quad \text{and} \tag{29}$$
$$\left[d - \frac{\partial}{\partial v} \log\left(\frac{d_{uv}}{d_u} - d\right)\right] \Big/ \left(\frac{d_{uv}}{d_u} - d\right).$$

Summarizing, we have this theorem.

THEOREM 12. *The velocity system $S(c, d)$ possesses the six absolute covariants*

$$
\begin{gathered}
\frac{c_v}{d_u}, \qquad \left(\frac{c_{uv}}{c_v} - c\right)\left(\frac{d_{uv}}{d_u} - d\right) \Big/ c_v, \\
\left[\frac{\partial}{\partial v} \log\left(\frac{c_{uv}}{c_v} - c\right)\right] \Big/ \left(\frac{d_{uv}}{d_u} - d\right), \\
\left[\frac{\partial}{\partial u} \log\left(\frac{d_{uv}}{d_u} - d\right)\right] \Big/ \left(\frac{c_{uv}}{c_v} - c\right), \\
\left[c - \frac{\partial}{\partial u} \log\left(\frac{c_{uv}}{c_v} - c\right)\right] \Big/ \left(\frac{c_{uv}}{c_v} - c\right), \\
\left[d - \frac{\partial}{\partial v} \log\left(\frac{d_{uv}}{d_u} - d\right)\right] \Big/ \left(\frac{d_{uv}}{d_u} - d\right).
\end{gathered} \tag{30}
$$

Any conformal covariant of a velocity system is a function of these and their conformal partial derivatives. These are defined by

$$
\frac{\partial}{\partial u} \Big/ \left(\frac{c_{uv}}{c_v} - c\right) \quad \textit{and} \quad \frac{\partial}{\partial v} \Big/ \left(\frac{d_{uv}}{d_u} - d\right). \tag{31}
$$

Let now $S(c, d)$ and $\Sigma(C, D)$ be such that there exist functions ϕ and ψ which satisfy the resulting covariant equations (30). If the operations on these six equations with the conformal derivatives (31) are reduced to identities by these functions ϕ and ψ, then $S(c, d)$ and $\Sigma(C, D)$ are conformally equivalent.

COLUMBIA UNIVERSITY AND
ILLINOIS INSTITUTE OF TECHNOLOGY

A COMPARISON OF ALGEBRAIC, METRIC, AND LATTICE BETWEENNESS

M. F. SMILEY

Introduction. We propose to investigate here the consequences of the identity of each pair chosen from three important generalizations of the relation of betweenness on a line, namely, algebraic betweenness [1, p. 27],[1] metric betweenness [3, p. 36], and lattice betweenness [7, Part II]. We shall also find an interpretation of metric betweenness in the Banach space of all continuous functions defined on the interval $0 \leqq t \leqq 1$ which can be used to establish the fact that this relation satisfies no strong four or five point transitivity [7, Part I] except t_1 and t_2.

We note first that algebraic betweenness implies metric betweenness and lattice betweenness. We find that algebraic betweenness and metric betweenness coincide in a seminormed real vector space[2] if and only if it is strictly convex in the sense of Clarkson [4, p. 404]. We then show that the coincidence of metric and lattice betweenness in a semimetric space [3, p. 38] which is also a lattice [2, p. 16] leads to a system which is a metric lattice (in the sense of G. Birkhoff [2, p. 41]). It follows that a complete seminormed real vector lattice is equivalent to an (L)-space [6] if and only if its metric and lattice betweenness relations are identical. Finally, we prove that algebraic and lattice betweenness coincide in a real vector lattice if and only if it is equivalent to the system of all real numbers. We conclude by giving the interpretation of metric betweenness in the space[3] $C[0, 1]$.

Presented to the Society, September 10, 1942; received by the editors March 27, 1942.

[1] References to the bibliography at the end of the paper will be in brackets.

[2] We shall use these terms as follows. A *seminormed real vector space* is a vector space over the field of all real numbers together with a real non-negative single-valued function $\|a\|$, called the "norm of a," satisfying (i) $\|\lambda a\| = |\lambda| \|a\|$, and (ii) $\|a\| = 0$ *if and only if* $a = 0$. A *normed real vector space* satisfies in addition (iii) $\|a\| + \|b\| \geqq \|a + b\|$. A *real vector lattice* is a vector space over the field of all real numbers which is also a lattice [2, p. 16] with respect to a partial ordering relation "$\geqq$" such that (i) $a \geqq b$ *and* $\lambda \geqq 0$ *implies* $\lambda a \geqq \lambda b$, and (ii) $a \geqq b$ *implies* $a + c \geqq b + c$ *for every* c. A *(semi)normed real vector lattice* is a real vector lattice which is also a (semi)normed real vector space; it is *complete* if evey fundamental sequence has a limit. A complete normed real vector space is usually called a (real) *Banach space.*

[3] The notation $C[0, 1]$ (sometimes simply C) is currently used to designate the space described in the concluding sentence of the preceding paragraph.

1. **The three betweenness relations.** We shall be interested in the following relations:

$$(a\ b\ c)_A \equiv b = \lambda a + (1-\lambda)c, \quad \lambda \text{ a real number, } 0 \leqq \lambda \leqq 1,$$
$$(a\ b\ c)_M \equiv \delta(a, b) + \delta(b, c) = \delta(a, c),$$
$$(a\ b\ c)_L \equiv (a \cap b) \cup (b \cap c) = b = (a \cup b) \cap (b \cup c),$$

to which we shall refer as R_A (algebraic betweenness), R_M (metric betweenness), and R_L (lattice betweenness). The relation R_A applies to a real vector space, R_M to a semimetric space, and R_L to a lattice. Our purpose is to find those systems to which each two of these three relations apply and have the same meaning.

The following three lemmas are known or evident.

LEMMA 1. *If S is a seminormed real vector space, then we have $(a\ b\ c)_M$ if and only if $\|a-b\|+\|b-c\|=\|a-c\|$.*

LEMMA 2. *If S is a distributive lattice, then $(a\ b\ c)_L$ if and only if $a \cap c \leqq b \leqq a \cup c$.*

PROOF. This is Lemma 9.1 of [7].

LEMMA 3. *If S is a seminormed real vector space, then $(a\ b\ c)_A \rightarrow (a\ b\ c)_M$ for every $a, b, c \in S$.*

LEMMA 4. *If S is a real vector lattice, then $(a\ b\ c)_A \rightarrow (a\ b\ c)_L$ for every $a, b, c \in S$.*

PROOF. It is well known that a vector lattice is distributive [2, p. 108], and hence by Lemma 2 it will suffice to show that $a \cap c \leqq b \leqq a \cup c$ is a consequence of $(a\ b\ c)_A$. We have the equations

$$(a \cup c) - b = (a-b) \cup (c-b) = (1-\lambda)(a-c) \cup -\lambda(a-c)$$
$$= ((a-c) \cup 0) - \lambda(a-c).$$

Note that $(a-c) \cup 0 \geqq \lambda((a-c) \cup 0)$, since $\lambda \leqq 1$; and that $\lambda((a-c) \cup 0) = \lambda(a-c) \cup 0 \geqq \lambda(a-c)$, since $\lambda \geqq 0$. Hence $(a \cup c) - b \geqq 0$, that is, $a \cup c \geqq b$. The other inequality is dual.

2. **Coincident metric and algebraic betweenness.** Clarkson [4, p. 404] calls a normed real vector space *strictly convex* if the equality $\|a\|+\|b\|=\|a+b\|$ for nonzero a and b implies that $a=\mu b$ for some $\mu>0$. The following theorem then tells us when metric and algebraic betweenness coincide.

THEOREM 1. *A seminormed real vector space S is strictly convex if and only if algebraic and metric betweenness coincide in S.*

PROOF. Let S be a strictly convex seminormed real vector space. By Lemma 3, we need only show that $(a\ b\ c)_M \rightarrow (a\ b\ c)_A$ for every $a, b, c \in S$. We may clearly suppose that $a-b$ and $b-c$ are nonzero. But then the strict convexity of S implies that $a-b=\mu(b-c)$ for some $\mu>0$. The choice $\lambda=1/1+\mu$ is effective for $(a\ b\ c)_A$. Conversely, if $(a\ b\ c)_M \rightarrow (a\ b\ c)_A$ for every $a, b, c \in S$, the equality $\|a\|+\|b\|=\|a+b\|$ for nonzero a and b implies $(a\ 0-b)_A$, and consequently that $0=\lambda a-(1-\lambda)b$ for real λ between 0 and 1. The choice $\mu=(1-\lambda)/\lambda$ is then effective in the definition of the strict convexity of S. The proof is complete.

REMARK 1. Clarkson [4, p. 413] also shows that every *separable* Banach space may be renormed to secure strict convexity.

3. **Coincident metric and lattice betweenness.** Our definition of lattice betweenness arose as a generalization of the metric betweenness of a metric lattice [2, p. 41; 5, 9]. We shall show in this section that, conversely, a semimetric space which is also a lattice has identical lattice and metric betweenness relations only if it is a metric lattice.[4] We apply this result to give a characterization of (L)-spaces in terms of betweenness relations.[5]

Let us consider a semimetric space M with distance function $\delta(a, b)$ which is also a lattice and in which lattice betweenness and metric betweenness coincide.

LEMMA 5. *If $a, b, c \in M$ and $a \leqq b \leqq c$, then $(a\ b\ c)_M$. If $a, b \in M$ then $(a\ a \cup b\ b)_M$ and $(a\ a \cap b\ b)_M$.*

PROOF. This is clear from Lemma 8.1 of [7].

LEMMA 6. *If $a, b \in M$, then $\delta(a, b)=\delta(a \cup b, a \cap b)$, and $\delta(a \cap b, b)=\delta(a, a \cup b)$.*

PROOF. Consider two elements $a, b \in M$. By Lemma 5 we have the relations $(a\ a \cup b\ b)_M$, $(a\ a \cap b\ b)_M$, $(a \cap b\ b\ a \cup b)_M$, and $(a \cap b\ a\ a \cup b)_M$. Hence we have the following equations:

---

[4] It should be noted that the results of §2 remain valid if we merely require the values of $\delta(a, b)$ to be in a real vector lattice, provided the terms *modular functional*, *metric lattice*, and so on, are given proper interpretations.

[5] G. Birkhoff's remark that $m[a] \equiv \|a_+\| - \|a_-\|$ is a sharply positive modular functional in an (L)-space (for proof, see [8]) aided the author materially in securing the present form of Theorem 2. The functional $m[a]$ is also *linear*. Its *homogeneity* follows readily from that of $\|a\|$, while its *additivity* is a simple consequence of the identity $m[a]-m[b]=\|a-b\|$ for $a \geqq b$. When the condition (IX) of Kakutani [6] holds, or, equivalently, when $m[a \cup b]-m[a \cap b]=\|a-b\|$, we find that $m[a]$ is the value of the Lebesgue-Stieltjes integral of the function corresponding to a.

$$\delta(a, a \cup b) + \delta(a \cup b, b) = \delta(a, b) = \delta(a, a \cap b) + \delta(a \cap b, b),$$
$$\delta(a \cap b, b) + \delta(b, a \cup b) = \delta(a \cap b, a \cup b) = \delta(a \cap b, a) + \delta(a, a \cup b).$$

Subtracting these equations we see that

$$\begin{aligned}\delta(a, a \cup b) - \delta(a \cap b, b) &= \delta(a, b) - \delta(a \cup b, a \cap b) \\ &= \delta(a \cap b, b) - \delta(a, a \cup b).\end{aligned}$$

It follows that the relations stated in the lemma hold.

REMARK 2. Note that the first result of Lemma 5 together with the results of Lemma 6 imply that the second result of Lemma 5 is valid. This is clear since $\delta(a, b) = \delta(a \cup b, a \cap b) = \delta(a \cup b, a) + \delta(a, a \cap b) = \delta(a, a \cap b) + \delta(a \cap b, b)$, and dually.

LEMMA 7. *The lattice M is modular.*

PROOF. We base our proof on the fact that if M fails to be modular then it contains as a sublattice the simplest non-modular lattice [2, p. 34]. Let the elements of this sublattice be O, a, b, c, I, with I the greatest, O the least element, and $b > c$. Since $O = a \cap c$ and $O < a < I$, Lemma 5 gives $\delta(O, c) = \delta(a, I)$. Since $O < c < b < I$, Lemma 5 and this result gives $\delta(a, I) + \delta(c, b) + \delta(b, I) = \delta(O, I)$. Again, since $I = a \cup b$, we have $\delta(O, I) + \delta(c, b) = \delta(O, I)$, and consequently $\delta(b, c) = 0$, contrary to $b > c$. Hence M cannot fail to be modular.

LEMMA 8. *If $a, b, c \in M$ and $a \geqq b$, then*

$$\delta(a \cup c, b \cup c) + \delta(a \cap c, b \cap c) = \delta(a, b).$$

PROOF. Consider elements $a, b, c \in M$ with $a \geqq b$. By Lemma 6 we have

$$\delta(a \cup c, b \cup c) = \delta(a \cup b \cup c, b \cup c) = \delta(a, a \cap (b \cup c)).$$

Noting that $a \geqq a \cap (b \cup c) \geqq b$, we find from Lemma 5 that

$$\delta(a \cup c, b \cup c) = \delta(a, b) - \delta(b, a \cap (b \cup c)).$$

Using Lemmas 7 and 6 we have

$$\begin{aligned}\delta(a \cup c, b \cup c) &= \delta(a, b) - \delta(b, b \cup (a \cap c)) \\ &= \delta(a, b) - \delta(a \cap c, b \cap a \cap c).\end{aligned}$$

Since $a \geqq b$, this proves the lemma.

LEMMA 9. *If $a^* \in M$ the functional $m[a] \equiv \delta(a \cup a^*, a^*) - \delta(a^*, a \cap a^*)$ is a sharply positive modular functional and $\delta(a, b) = m[a \cup b] - m[a \cap b]$ for every $a, b \in M$.*

PROOF. From Lemma 6 and the fact that $\delta(a, b)>0$ for $a\neq b$, it will suffice to show that

$$\delta(a, b) = m[a] - m[b] \tag{1}$$

for every $a, b\in M$ such that $a\geqq b$. Hence consider elements $a, b\in M$ for which $a\geqq b$. Lemma 8 yields the equation

$$\delta(a\cup a^*, b\cup a^*) + \delta(a\cap a^*, b\cap a^*) = \delta(a, b). \tag{2}$$

Since $a\cup a^*\geqq b\cup a^*\geqq a^*$, and dually, Lemma 5 yields the following equations

$$\delta(a\cup a^*, a^*) - \delta(b\cup a^*, a^*) = \delta(a\cup a^*, b\cup a^*),$$
$$\delta(a^*, b\cap a^*) - \delta(a^*, a\cap a^*) = \delta(a\cap a^*, b\cap a^*).$$

Adding these equations and using the equation (2) we obtain (1) from the definition of the functional $m[a]$. The proof is complete.

REMARK 3. The validity of Lemmas 7–9 depends only on Lemma 6 and the first result of Lemma 5.

LEMMA 10. *The distance function $\delta(a, b)$ is a metric for M.*

PROOF. This follows from Lemma 9 and Theorem 3.10 of [2].

THEOREM 2. *If S is a semimetric space with distance function $\rho(a, b)$ which is also a lattice, then metric betweenness and lattice betweenness coincide in S if and only if:*

(i) *For every $a, b, c\in S$ the inequalities $a\leqq b\leqq c$ imply that $(a\ b\ c)_M$.*

(ii) *For every $a, b\in S$, $\rho(a, b)=\rho(a\cup b, a\cap b)$ and $\rho(a, a\cup b)=\rho(b, a\cap b)$.*

The conditions (i) *and* (ii) *hold if and only if for each $a^*\in M$ the functional $m[a]\equiv\rho(a\cup a^*, a^*)-\rho(a^*, a\cap a^*)$ is a sharply positive modular functional and S is a metric lattice with metric $\rho(a, b)=m[a\cup b]-m[a\cap b]$.*

PROOF. This is clear from Lemmas 5–9, Remarks 2 and 3, and Theorem 10.1 of [7].

We also have a specialization of Theorem 2.

THEOREM 3. *A complete seminormed real vector lattice S is equivalent to an (L)-space if and only if lattice and metric betweenness coincide in S.*

PROOF. Let S be a complete seminormed real vector lattice in which lattice and metric betweenness coincide. By Theorem 2, Lemma 6, and Lemma 10, S is a complete normed real vector lattice satisfying the conditions,

(VIII) if $a, b \in S$ and $a \geqq 0$, $b \geqq 0$, then $\|a\| + \|b\| = \|a+b\|$;

(IX) if $a, b \in S$ and $a \cap b = 0$, then $\|a+b\| = \|a-b\|$;

of Kakutani [6]. Kakutani's fundamental Theorem 7 [6] together with the fact that his condition (V) is not needed (in his Theorem 7) [8] now shows that S is equivalent to an (L)-space.

Conversely, if S is equivalent to an (L)-space, it is easy to verify the conditions (i) and (ii) of Theorem 2. Hence lattice betweenness and metric betweenness coincide in S. The proof is complete.

REMARK 4. We may restate Theorem 3 as follows. *A complete semi-normed real vector lattice is equivalent to an (L)-space if and only if it is a metric lattice.*

4. Coincident lattice and algebraic betweenness. We include the following theorem for the sake of completeness.

THEOREM 4. *If S is a real vector lattice, then lattice betweenness and algebraic betweenness coincide in S if and only if the dimension of S (as a vector space) is one.*

PROOF. We need only show that $(a\ b\ c)_L \rightarrow (a\ b\ c)_A$ for every $a, b, c \in S$ implies that S is one-dimensional. It is easy to verify that R_A has the fundamental transitivity t_3 [7]. By Theorem 9.7 of [7] every pair of elements of S are comparable by "$\geqq$." If, then, there were two linearly independent elements $a, b \in S$, we should have (by symmetry) $a > b > 0$ or $a > 0 > b$; and hence [7, Lemma 8.1] $(a\ b\ 0)_L$ or $(a\ 0\ b)_L$. By hypothesis, we would have $b = \lambda a$ or $0 = \lambda a + (1-\lambda)b$, which clearly contradicts the linear independence of a and b. Thus S is one-dimensional.

5. Metric betweenness in the space $C[0, 1]$. The question of whether or not the metric betweenness of a normed real vector lattice possesses any more than the required transitivities (that is, (α), (β), t_1, and t_2 of [7]) can be settled by the space $C[0, 1]$. We give in this section an interpretation of R_M in $C[0, 1]$ and leave to the reader the construction of simple examples which show that this relation has only the required transitivities. This result should be contrasted with the result of Clarkson mentioned in Remark 1.

LEMMA 11. *In the space $C[0, 1]$ the relation $(a\ 0\ b)_M$ holds if and only if there is a point t_0 on $[0, 1]$ at which each of the functions $|a(t)|$, $|b(t)|$, and $|a(t)-b(t)|$ attains its maximum value and the relation $(a(t_0)\ 0\ b(t_0))_M$ holds.*

PROOF. Suppose that the relation $(a\ 0\ b)_M$ holds in $C[0, 1]$. We may clearly assume that a and b are nonzero. Let t_0 be a point on $[0, 1]$ at

which the function $|a(t)-b(t)|$ attains its maximum value. Then we have

$$\|a\| + \|b\| = \|a - b\| = |a(t_0) - b(t_0)|$$
$$\leqq |a(t_0)| + |b(t_0)| \leqq \|a\| + \|b\|.$$

It follows that $|a(t_0)| + |b(t_0)| = \|a\| + \|b\|$ and that each of the functions $|a(t)|$ and $|b(t)|$ attains its maximum at t_0. Consequently neither $a(t_0)$ nor $b(t_0)$ can be zero. If they should both have the same sign we would find that $|a(t_0)-b(t_0)| < |a(t_0)| + |b(t_0)|$. This is impossible, and we conclude that the relation $(a(t_0)\ 0\ b(t_0))_M$ holds.

Conversely, if there is a point t_0 on $[0, 1]$ at which each of the functions $|a(t)|$, $|b(t)|$, and $|a(t)-b(t)|$ attains its maximum value and the relation $(a(t_0)\ 0\ b(t_0))_M$ holds, we see that

$$\|a - b\| = |a(t_0) - b(t_0)| = |a(t_0)| + |b(t_0)| = \|a\| + \|b\|.$$

The proof is complete.

THEOREM 5. *In the space $C[0, 1]$ the relation $(a\ b\ c)_M$ holds if and only if there is a point t_0 on $[0, 1]$ at which each of the functions $|a(t)-b(t)|$, $|b(t)-c(t)|$, and $|a(t)-c(t)|$ attains its maximum value and the relation $(a(t_0)\ b(t_0)\ c(t_0))_M$ holds.*

PROOF. This is an immediate consequence of Lemma 11 and the fact that $(a\ b\ c)_M$ if and only if $(a+x\ b+x\ c+x)_M$ for every $a, b, c, x \in C[0, 1]$.

BIBLIOGRAPHY

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.

2. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications vol. 25, 1940.

3. L. M. Blumenthal, *Distance geometries*, University of Missouri Studies vol. 12 (1938) no. 2.

4. J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. vol. 40 (1936) pp. 396–414.

5. V. Glivenko, *Géométrie des systèmes de choses normées*, Amer. J. Math. vol. 58 (1936) pp. 799–828.

6. S. Kakutani, *Concrete representation of abstract (L)-spaces and the mean ergodic theorem*, Ann. of Math. (2) vol. 42 (1941) pp. 523–537.

7. Everett Pitcher and M. F. Smiley, *Transitivities of betweenness*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 95–114.

8. M. F. Smiley, *A remark on S. Kakutani's characterization of (L)-spaces*, Ann. of Math. (2) vol. 43 (1942) pp. 528–529.

9. L. R. Wilcox and M. F. Smiley, *Metric lattices*, Ann. of Math. (2) vol. 40 (1939) pp. 309–327.

LEHIGH UNIVERSITY

ABSTRACT POLYNOMIALS IN NON-ABELIAN GROUPS

KNOX MILLSAPS

Introduction. The aim of this note is to give some generalizations for groups of the theories of abstract polynomials as developed by Fréchet, Gateaux, Martin, Mazur, Michal, Orlicz and more recently, Van der Lijn.[1] Although the theories are equivalent for functions with arguments and values in abelian groups,[2] this equivalence is not the case when the argument and value groups are non-abelian.[3]

In §1, a calculus of finite differences for functions with arguments and values in non-abelian groups which contain no elements of finite order is defined, and the fundamental definitions of polynomials and monomials are stated. In §2, the homogeneity in the increment of the n-difference of a polynomial of degree n is proved, and the theorem on unique pseudo-decomposition is proved after giving some preliminary theorems on the structure of differences of arbitrary functions and polynomials. In §3, a brief discussion of the extensions to non-abelian groups as value spaces is given.

I should like to thank Professor A. D. Michal for his helpful suggestions and constructive criticisms during the preparation of this note.

1. **Definitions.** For the purposes of polynomial theory the value groups are restricted to contain no elements of finite order.

To construct the calculus of finite differences for functions with arguments and values in non-abelian groups, we define

$$
{}_{\omega}\Delta[\alpha_1;\beta_1]f(x) \equiv \begin{cases} f(x\omega)f^{-1}(x) & \text{if } \alpha_1 = 1 \text{ and } \beta_1 = 1, \\ f(\omega x)f^{-1}(x) & \text{if } \alpha_1 = 1 \text{ and } \beta_1 = 2, \\ f^{-1}(x)f(x\omega) & \text{if } \alpha_1 = 2 \text{ and } \beta_1 = 1, \\ f^{-1}(x)f(\omega x) & \text{if } \alpha_1 = 2 \text{ and } \beta_1 = 2, \end{cases}
$$

Presented to the Society, April 11, 1942; received by the editors April 21, 1942.

[1] References to these theories will be found in the bibliography. This list will be referred to by numbers in brackets.

[2] The equivalence of some of these definitions was proved by Martin in his California Institute of Technology thesis, 1932, and of the remaining definitions by Van der Lijn. A summary has been given by Van der Lijn [1, pp. 78–80].

[3] If the abstraction of additivity is multiplicative distributivity, then the generalizations of the definitions of Mazur and Orlicz [1, p. 63] and Van der Lijn, [1, pp. 60–61] are not equivalent; this is easily seen by considering $f(x) \equiv x^k$.

and inductively

$$
\begin{aligned}
&{}_{\omega}^{n}\Delta[\alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n]f(x) \\
&\qquad \equiv {}_{\omega}\Delta[\alpha_n; \beta_n]\,{}_{\omega}^{n-1}\Delta[\alpha_1, \cdots, \alpha_{n-1}; \beta_1, \cdots, \beta_{n-1}]f(x).
\end{aligned}
$$

DEFINITION OF A MONOMIAL. *A* $[\alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n]$ *monomial*[4] *is a function* $f(x)$ *which satisfies for all* x *and* ω *the following functional equation*

$$
{}_{\omega}^{n}\Delta[\alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n]f(x) = f^{n!}(\omega).
$$

The obviously unique value of n is termed the degree of the monomial.

DEFINITION OF A POLYNOMIAL. *A function* $f(x)$ *which satisfies the following functional equation for all* x *and* ω

$$
{}_{\omega}^{n+1}\Delta[\alpha_1, \cdots, \alpha_{n+1}; \beta_1, \cdots, \beta_{n+1}]f(x) = 1
$$

will be called a $[\alpha_1, \cdots, \alpha_{n+1}; \beta_1, \cdots, \beta_{n+1}]$ *polynomial.*[5]

The least value of n for which the above equation holds will be termed the degree of the polynomial.

The binomial coefficients are denoted in the usual manner; $\Gamma_{k,n}^{i}$ is defined by

$$
\Gamma_{k,n}^{i} \equiv \sum \frac{n!}{\alpha_1!\alpha_2! \cdots \alpha_{k+1}!}
$$

where $\sum_{i=1}^{k+1}\alpha_i = n$, $\sum_{i=2}^{k+1}(q-1)\alpha_q = i$, $i \neq nk$, and $i \neq 0$. An arbitrary element of the central subgroup will be denoted by ϑ. To simplify notation, the dropping of unnecessary indices implies that the value of a product of factors is independent of the order of the particular factors controlled by the dropped indices.

The next equation is a generalization of an identity due to Marchaud[6]

$$
{}_{\omega}^{k+1}\Delta[\beta_1, \cdots, \beta_n]f(x) = \sum_{i=0}^{nk}\sum_{j}{}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]f[P_{ij}],
$$

---

[4] The inner automorphisms of a group are interesting examples of a $[1; 2]$ monomial; similarly, the canonical transformations of quantum mechanics.

[5] If the elements of a group are taken to be the $n \times n$ matrices whose elements are in a commutative field and whose determinants do not vanish, and if the group operation is defined as row by column multiplication, then $f[\|x_j^i\|] = \|a_k^i x_m^k x_n^m b_j^n\|$ is a $[1, 1; 2, 1]$ polynomial.

[6] Marchaud [1, p. 368].

where $j=1, \cdots, \Gamma_{k,n}^{i}$ and P_{ij} denotes a particular permutation of $\omega, \omega, \cdots, \omega$ (i times) and x.

2. **Fundamental theorems.** Let $A(x)$ denote a function with arguments in a non-abelian group and values in an abelian group. The theorems of Van der Lijn for functions with arguments and values in abelian groups can be extended with a few immediate changes to hold for differences of the type ${}_{\omega}^{n}\Delta A(x)$. This section is devoted to abstractions for differences of the kind ${}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x)$, where ω is an arbitrary element.

THEOREM I. *If $A(x)$ is a $[\beta_1, \cdots, \beta_{n+1}]$ polynomial of degree n, and if k is an integer, then*

$$ {}_{\omega^k}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) = k^n \, {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x). $$

PROOF. For $k=0$ or 1, the theorem is trivial. For $k>1$, we hypothetically have

$$ {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x, \omega^{i}) - {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x, \omega^{i-1}) = 0. $$

By a few manipulations and Marchaud's identity we get

$$ \begin{aligned} {}_{\omega^k}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) &= \sum_{i=0}^{n(k-1)} \Gamma_{k-1,n}^{i} \, {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) \\ &= k^n \, {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x). \end{aligned} $$

For $k<0$, we evidently have

$$ \begin{aligned} {}_{\omega^k}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) &= {}_{(\omega^{-1})^{-k}}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) \\ &= (-k)^n \, {}_{\omega^{-1}}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) \\ &= (-)^n(-k)^n \, {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) \\ &= k^n \, {}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x). \end{aligned} $$

THEOREM II. *If ${}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x) = g(\omega)$, where $g(\omega)$ is independent of x, then $g(\omega)$ is a $[\beta_1, \cdots, \beta_n]$ monomial of degree n in ω.*

THEOREM III. *If $A(x)$ is a $[\beta_1, \cdots, \beta_{n+1}]$ polynomial of degree n, then ${}_{\omega}^{n}\Delta[\beta_1, \cdots, \beta_n]A(x)$ is independent of x.*

PROOF. If $n+2$ elements of the argument group are denoted by x_i, where

$$ x_{i+1} = \prod(x_i, \xi), \qquad i = 0, \cdots, n+1, $$

then by steps roughly analogous to the usual proof of similar theorems we can derive

$$\sum_{p=0}^{n}\sum_{i=0}^{n+1}C_{n+1,i}C_{n,p}(-)^{i}i^{n-p}k^{p}\,{}_{\omega}^{n}\Delta[\beta_1,\cdots,\beta_n]A(x_i)=0.$$

In the last equation we see that $\sum_{p=0}^{n}C_{n,p}k^p$ is a simple polynomial. From this observation we deduce[7]

$$\sum_{i=0}^{n+1}(-)^{i}C_{n+1,i}i^{n-p}\,{}_{\omega}^{n}\Delta[\beta_1,\cdots,\beta_n]A(x_i)=0,\qquad p=0\cdots,n.$$

If we consider one difference as given and the remaining $n+1$ differences as unknown, then the unique solutions are

$${}_{\omega}^{n}\Delta[\beta_1,\cdots,\beta_n]A(x_i)={}_{\omega}^{n}\Delta[\beta_1,\cdots,\beta_n]A(x_0),\qquad i=1,\cdots,n+1,$$

for the determinant of the system of equations is given by

$$(-)^{\delta(n)}\prod_{i=1}^{n+1}C_{n+1,i}V\neq 0$$

where V is a Vandermonde determinant and $\delta(n)=1$, if $n\equiv 1, 2 \bmod 4$; $\delta(n)=2$, if $n\equiv 0, 3 \bmod 4$.

THEOREM IV. *If $A(x)$ is a $[\beta_1,\cdots,\beta_{n+1}]$ polynomial of degree n, and if $\rho_n=\prod_{i=0}^{n-1}(n-i)!$, then we have the unique pseudo-decomposition*

$$\rho_n A(x)=\sum_{i=0}^{n}M_i(x),$$

where $M_i(x)$ are $[\beta_1,\cdots,\beta_i]$ monomials of degree i.

3. **Further discussion.** With the generalized definition of a polynomial an extension of Theorem I for functions with arguments and values in non-abelian groups can be made. The trivial converse of Theorem III can be proved in an obvious manner. The non-abelian analogue of Theorem IV does not hold, for counterexamples can be exhibited.[8]

In conclusion, it is remarked that the difficulty with the non-abelian valued case is an inability to solve explicitly a system of group equations.

---

[7] It can be shown that if a simple polynomial $\sum_{i=0}^{n}a_ix^i$ of degree n vanishes for $n+1$ distinct values of x, then the coefficients vanish.

[8] With the group defined in (5), a trivial, but sufficient, example is $\|x_j^i\|=\|x_i^i\|\,\|\delta_j^i\|=\|\delta_j^i\|\,\|x_j^i\|$.

BIBLIOGRAPHY

M. FRÉCHET

1. *Une définition fonctionnelle des polynômes*, Nouvelles Annales de Mathématiques (4) vol. 9 (1909) pp. 145–162.
2. *Les polynômes abstraits*, Journal de Mathématiques vol. 8 (1929) pp. 71–92.

R. GATEAUX

1. *Sur diverses questions du calcul fonctionnel*, Bull. Soc. Math. France vol. 50 (1922) pp. 1–37.

A. MARCHAUD

1. *Sur les dérivées et sur les différences des fonctions de variables réelles*, Journal de Mathématiques vol. 6 (1927) pp. 337–425.

S. MAZUR AND W. ORLICZ

1. *Grundlegende Eigenschaften der polynomische Operationen*, Studia Mathematica vol. 5 (1934) pp. 58–68, 179–189.

A. D. MICHAL AND R. S. MARTIN

1. *Some expansions in vector space*, Journal de Mathématiques vol. 13 (1934) pp. 69–91.

G. VAN DER LIJN

1. *Les polynômes abstraits*, Bull. Sci. Math. vol. 64 (1940) pp. 55–80, 102–112, 163–196.

CALIFORNIA INSTITUTE OF TECHNOLOGY

ON THE QUADRIC OF LIE

SU-CHENG CHANG

The most important quadric attached to an ordinary point of a non-ruled and nondegenerate surface is, perhaps, the quadric of Lie. The characteristic curve of the quadric of Lie varying along an asymptotic curve of the surface decomposes into an asymptotic tangent and two edges of the quadrilateral of Demoulin.[1] In this note we propose to determine whether the characteristic curve of the quadric of Lie may decompose into two conics when the quadric of Lie varies along certain curves of the surface. The answer is positive.

Let (u, v) be the asymptotic net of a surface (M) and (M, M_1, M_2, M_3) its normal tetrahedron of Cartan, MM_1, MM_2 being the two asymptotic tangents and MM_3, M_1M_2 being the directrices of Wilczynski. Except for a projective transformation the surface (M) is determined by the system

Received by the editors June 8, 1942.

[1] A. Demoulin, *Sur la théorie des lignes asymptotiques*, C. R. Acad. Sci. Paris vol. 147 (1908) pp. 413–415, *Sur la quadrique de Lie*, ibid. pp. 493–496, *Sur quelques propriétés des surfaces courbes*, ibid. pp. 565–568.

$$
\begin{aligned}
&\frac{\partial M}{\partial u} = (1/2)\frac{\partial \log \gamma}{\partial u} M + M_1,\\
&\frac{\partial M_1}{\partial u} = B^2 M - (1/2)\frac{\partial \log \gamma}{\partial u} M_1 + \beta M_2,\\
&\frac{\partial M_2}{\partial u} = KM + (1/2)\frac{\partial \log \gamma}{\partial u} M_2 + M_3,\\
&\frac{\partial M_3}{\partial u} = A^2\beta M + KM_1 + B^2 M_2 - (1/2)\frac{\partial \log \gamma}{\partial u} M_3,\\
&\frac{\partial M}{\partial v} = (1/2)\frac{\partial \log \beta}{\partial v} M + M_2,\\
&\frac{\partial M_1}{\partial v} = \overline{K} M + (1/2)\frac{\partial \log \beta}{\partial v} M_1 + M_3,\\
&\frac{\partial M_2}{\partial v} = A^2 M - (1/2)\frac{\partial \log \beta}{\partial v} M_2 + \gamma M_1,\\
&\frac{\partial M_3}{\partial v} = B^2\gamma M + A^2 M_1 + \overline{K} M_2 - (1/2)\frac{\partial \log \beta}{\partial v} M_3,
\end{aligned}
\tag{1}
$$

where M denotes a point of the surface with the coordinates M^i $(i=1, 2, 3, 4)$ and

$$2K = \beta\gamma - \frac{\partial^2 \log \beta}{\partial u \partial v}, \qquad 2\overline{K} = \beta\gamma - \frac{\partial^2 \log \gamma}{\partial u \partial v}.$$

The conditions of integrability of the system (1) are found to be

$$
\begin{aligned}
&\frac{\partial A^2}{\partial u} = K\frac{\partial(\log \beta K)}{\partial v}, \qquad \frac{\partial B^2}{\partial v} = \overline{K}\frac{\partial(\log \gamma \overline{K})}{\partial u},\\
&A\frac{\partial(A\beta)}{\partial v} = B\frac{\partial(B\gamma)}{\partial u}.
\end{aligned}
\tag{2}
$$

As the coordinates of any point in space can be expressed in the form

$$P = y_1 M + y_2 M_1 + y_3 M_2 + y_4 M_3, \tag{3}$$

(y_1, y_2, y_3, y_4) being the local coordinates of P with respect to the tetrahedron $\{M, M_1, M_2, M_3\}$, we have the equation of the quadric of Lie at M, namely,

$$y_1 y_4 - y_2 y_3 = 0. \tag{4}$$

Suppose that the point P be fixed in space, from (1) and (3) it follows that

$$
\begin{aligned}
\frac{\partial y_1}{\partial u} &= -(1/2)\frac{\partial \log \gamma}{\partial u} y_1 - B^2 y_2 - K y_3 - A^2\beta y_4, \\
\frac{\partial y_2}{\partial u} &= -y_1 + (1/2)\frac{\partial \log \gamma}{\partial u} y_2 - K y_4, \\
\frac{\partial y_3}{\partial u} &= -\beta y_2 - (1/2)\frac{\partial \log \gamma}{\partial u} y_3 - B^2 y_4, \\
\frac{\partial y_4}{\partial u} &= -y_3 + (1/2)\frac{\partial \log \gamma}{\partial u} y_4, \\
\frac{\partial y_1}{\partial v} &= -(1/2)\frac{\partial \log \beta}{\partial v} y_1 - \overline{K} y_2 - A^2 y_3 - B^2\gamma y_4, \\
\frac{\partial y_2}{\partial v} &= -(1/2)\frac{\partial \log \beta}{\partial v} y_2 - \gamma y_3 - A^2 y_4, \\
\frac{\partial y_3}{\partial v} &= -y_1 + (1/2)\frac{\partial \log \beta}{\partial v} y_3 - \overline{K} y_4, \\
\frac{\partial y_4}{\partial v} &= -y_2 + (1/2)\frac{\partial \log \beta}{\partial v} y_4.
\end{aligned}
\tag{5}
$$

Differentiating (4) along a curve $v=v(u)$ of the surface (M) and making use of (5), we obtain that the characteristic of the quadric of Lie along this curve is given by (4) and

$$
\beta(y_2^2 - A^2 y_4^2) + \frac{dv}{du}\gamma(y_3^2 - B^2 y_4^2) = 0. \tag{6}
$$

The latter represents two planes when and only when

$$
\frac{dv}{du} = -\frac{\beta A^2}{\gamma B^2}. \tag{7}
$$

Therefore *there are curves L of a one-parameter system on a surface along each of which two consecutive quadrics of Lie intersect in two conics C_1 and C_2.* It is easily seen that the planes of these conics at M are

$$
By_2 - Ay_3 = 0 \tag{8$_1$}
$$

and

$$(8_2) \qquad By_2 + Ay_3 = 0,$$

respectively.

In order that *the curves L should be indeterminate it is necessary and sufficient that the surface in consideration is of Demoulin and Godeaux.*[2]

In general *the conics C_1 and C_2 have M and M_3 in common.*

Noticing that the vertices of the quadrilateral of Demoulin are

$$(9) \qquad D_{\epsilon\epsilon'} = \epsilon\epsilon' ABM + \epsilon AM_1 + \epsilon' BM_2 + M_3,$$

where $\epsilon = \pm 1$, $\epsilon' = \pm 1$ and that the planes (8_1) and (8_2) pass through the two diagonals of the quadrilateral of Demoulin, respectively, we infer that *the two conics C_1 and C_2 can be constructed by means of the quadric of Lie and the quadrilateral of Demoulin.*

The second quadric Φ_1, namely, the associate quadric,[3] in the sequence of Godeaux[4] originally defined in virtue of the representation in S_8, is given by the equation

$$(10) \qquad y_1^2 - B^2y_2^2 - A^2y_3^2 + A^2B^2y_4^2 + \frac{N}{\beta\gamma}(y_1y_4 - y_2y_3) = 0$$

where

$$N = B(B\gamma)_u = A(A\beta)_v.$$

It is obvious that *each of the planes (8_i) $(i=1, 2)$ intersects the associate quadric in a conic which is tangent to the conic C_i $(i=1, 2)$ at two opposite vertices of the quadrilateral of Demoulin.*

When the point M varies along a curve L of the system the conics C_1, C_2 and their consecutive conics always lie on the quadric of Lie at M.

In fact, a point on the conic C_1 at M can be given parametrically in the form $(AB\rho^2, A\rho, B\rho, 1)$, where ρ is a parameter. By means of (5) we can expand the local coordinates of the points on the consecu-

---

[2] L. Godeaux, *Sur une classe de surfaces*, Bulletin de l'Académie royale de Belgique vol. 18 (1932) pp. 1015–1025.

[3] B. Su, *On the surfaces whose asymptotic curves belong to linear complexes*, I, Tôhoku Math. J. vol. 40 (1935) pp. 408–420; II, ibid. pp. 433–448; III, ibid. vol. 41 (1935) pp. 1–9; IV, ibid. pp. 203–215; V, Science Reports of the Tôhoku Imperial University vol. 25 (1935) pp. 601–633; VI, ibid. pp. 634–642; *On the surface whose Lie quadrics all touch a fixed plane*, Science Reports of the National University of Chekiang vol. 2 (1936) pp. 39–51. See also L. Godeaux, *Remarques sur les quadriques associées aux points d'une surface*, Journal of the Chinese Mathematical Society vol. 2 (1937) pp. 1–5.

[4] L. Godeaux, *La théorie des surfaces et l'espace réglé*, Paris, 1934. A new definition of the sequence of Godeaux is given by the present author. See S. C. Chang, *Some theorems on ruled surfaces*, appear in Science Record, Academia Sinica vol. 1 (1942).

tive conic of C_1 into the following series:

$$
\begin{aligned}
y_1 &= AB\rho^2 + \Big[-(1/2)\Big(\frac{\partial \log \gamma}{\partial u} - \frac{\beta A^2}{\gamma B^2}\frac{\partial \log \beta}{\partial v}\Big)AB\rho^2 - AB^2\rho \\
&\quad - KB\rho - \frac{\beta A^2}{\gamma B^2}(\overline{K}A\rho + A^2B\rho)\Big]\Delta u + (\Delta u^2), \\
y_2 &= A\rho + \Big[-AB\rho^2 + (1/2)\frac{\partial \log \gamma}{\partial u}A\rho - K \\
&\quad - \frac{\beta A^2}{\gamma B^2}\Big(-(1/2)\frac{\partial \log \beta}{\partial v}A\rho - B\gamma\rho - A^2\Big)\Big]\Delta u + (\Delta u^2), \\
y_3 &= B\rho + \Big[-\beta A\rho - (1/2)\frac{\partial \log \gamma}{\partial u}B\rho - B^2 \\
&\quad - \frac{\beta A^2}{\gamma B^2}\Big(-AB\rho^2 + (1/2)\frac{\partial \log \beta}{\partial v}B\rho - \overline{K}\Big)\Big]\Delta u + (\Delta u^2), \\
y_4 &= 1 + \Big[-B\rho + (1/2)\frac{\partial \log \gamma}{\partial u} \\
&\quad - \frac{\beta A^2}{\gamma B^2}\Big(-A\rho + (1/2)\frac{\partial \log \beta}{\partial v}\Big)\Big]\Delta u + (\Delta u^2).
\end{aligned}
\tag{11}
$$

Substituting (11) into (4) gives that the coefficients of the terms $(\Delta u)^0$ and $(\Delta u)^1$ vanish for any value of ρ. A similar result holds for the conic C_2, which completes the proof.

NATIONAL UNIVERSITY OF CHEKIANG

TRANSFORMATIONS OF MULTIPLE FOURIER SERIES

L. B. HEDGE

1. **Introduction.** The object of the present paper is the study and characterization of certain classes of factor sequence transformations of multiple Fourier series. A recent moment problem solution[1] by the author and a scheme of summation of multiple Fourier series developed by Bochner[2] are used in the study. The results include and extend known results for single Fourier series.

2. **Definitions and notation.** Let n be a positive integer, fixed but arbitrary. R^n will denote the euclidean n-space. (x), (y), and so on will denote $(x_1, x_2, \cdots, x_n)$, $(y_1, y_2, \cdots, y_n)$, and so on, points of R^n. ν, τ, j, k, s will be used for non-negative integers, and (ν), (τ), and so on will be used for $(\nu_1, \nu_2, \cdots, \nu_n)$, $(\tau_1, \tau_2, \cdots, \tau_n)$, lattice points of R^n. (0) will mean $(0, 0, \cdots, 0)$, and $(x)=(y)$ will mean $x_j=y_j$, $j=1, 2, \cdots, n$. $(k\cdot x)$ will stand for the number $k_1x_1+k_2x_2+\cdots+k_nx_n$, $|x|$ for the number $(x_1^2+x_2^2+\cdots+x_n^2)^{1/2}$. A, I, and λ will be used for functions defined on the lattice points of R^n, and I will be the characteristic function of the lattice points of R^n. E will be the set $E_{(x)}(-\pi\leqq x_j<\pi, j=1, 2, \cdots, n)$. R and t will be used for real numbers. $(x+y)$ will stand for $(x_1+y_1, x_2+y_2, \cdots, x_n+y_n)$, and $B(n)$ for a real constant depending only on n. Φ will be used for a function U^* of bounded variation in the sense of Saks, and if $\Phi(H)=\Phi_1(H)+\Phi_2(H)$ for any Borel set H with $\Phi_1(H)\geqq 0\geqq\Phi_2(H)$ we will write $\int_H f(x)\,|d\Phi(E)|$ for $\int_H f(x)d\Phi_1(E)-\int_H f(x)d\Phi_2(E)$. When Φ is the Lebesgue measure function we will write $\int_H f(x)dx$ for

$$\int_H f(x)d\Phi(E).$$

We will write $f\in L$ to indicate that $\int_E f(x)dx$ exists, and $f\in C$ to indicate that f is continuous on $\overline{E}$ and $f(x)=f(x+y)$ for all combinations of $y_j=0$ or 2π, $j=1, 2, \cdots, n$. A function f defined over E will be defined over R_n by the extension $f(x)=f(x+y)$ with $y_j=0$ or 2π, $j=1, 2, \cdots, n$.

Let

Presented to the Society, December 30, 1940; received by the editors May 17, 1942.

[1] L. B. Hedge, *Moment problem for a bounded region*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 282–285. Referred to later as MP.

[2] S. Bochner, *Summation of multiple Fourier series by spherical means*, Trans. Amer. Math. Soc. vol. 40 (1936) pp. 175–207.

$$\mathfrak{S}(\lambda, A, x) \simeq \sum_{(k)} \lambda(k)A(k)e^{i(k\cdot x)}, \tag{1}$$

and

$$\mathfrak{S}_\nu(\lambda, A, x) = \sum_{|k|^2 \leq \nu} \lambda(k)A(k)e^{i(k\cdot x)}. \tag{2}$$

If

$$\lambda(k)A(k) = (2\pi)^{-n} \int_E f(x)e^{-i(k\cdot x)}dx \tag{3}$$

or

$$\lambda(k)A(k) = (2\pi)^{-n} \int_E e^{-i(k\cdot x)}d\Phi(E), \qquad \int_E |d\Phi(E)| < \infty, \tag{4}$$

we will write $\mathfrak{S}(f, x)$ or $\mathfrak{S}(d\Phi, x)$, respectively, for the left side of equation (1), and similarly alter equations (2) and (5).

We write

$$S_R(\lambda, A, x) = \sum_{(k)} \Psi(|k|/R)\lambda(k)A(k)e^{i(k\cdot x)}, \tag{5}$$

where

$$\Psi(t) = e^{-t^2},$$

for the Bochner-Gauss[2] spherical means of the sequence (2), and

$$K_R(u) = \sum_{(k)} \Psi\left(\frac{|K|}{R}\right) e^{-i(k\cdot u)} = S_R(I, I, u) \tag{6}$$

for the corresponding kernel.

3. **Spherical summation.** We proceed to some modification (largely notational) of the Bochner summation theory. The transformation which takes (2) into (5) is given by the matrix $T: \|a_{R,\nu}\|$, where $a_{R,\nu} = \Psi(R_\nu/R) - \Psi(R_{\nu+1}/R)$, and $\{R_\nu\}$ is a subsequence of $\{0, 1, 2^{1/2}, 3^{1/2}, 4^{1/2}, \cdots, k^{1/2}, \cdots\}$. We have at once

$$a_{R,\nu} \geqq 0, \qquad \lim_{R\to\infty} a_{R,\nu} = 0, \qquad \sum_{\nu=0}^{\infty} a_{R,\nu} = 1,$$

whence T is a regular Toeplitz transformation.[3]

---

[3] See, for example, A. Zygmund, *Trigonometrical series*, Monografje Matematyczne, vol. 5, Warsaw, 1935, pp. 79–86. That R is a continuous variable is unimportant.

The following properties of the summation scheme follow immediately from Bochner's work:

$$0 \leqq K_R(u) \leqq M_R < \infty, \tag{7}$$

$$S_R(f, x) = B(n) \int_E f(u) K_R(u - x) du, \qquad \text{if } f \in L, \tag{8}$$

and,

$$S_R(f, x) \to f(x) \tag{9}$$

at every point (x) of continuity of f, and uniformly on E if $f \in C$.

4. **Classes of multiple Fourier series.** The theorem of MP may now be given in the following form:

THEOREM. *In order that* $\mathfrak{S}(\lambda, A, x) = \mathfrak{S}(d\Phi, x)$ *with*

$$1. \int_E |d\Phi| \leqq M, \quad \text{or} \quad 2. \ \Phi \geqq 0,$$

it is necessary and sufficient that

$$1. \int_E |S_R(\lambda, A, x)| \, dx \leqq M, \quad \text{or} \quad 2. \ S_R(\lambda, A, x) \geqq 0,$$

and in order that $\mathfrak{S}(\lambda, A, x) = \mathfrak{S}(f, x)$ *with*

$$3. \ f \in L, \quad \text{or} \quad 4. \ |f| \leqq M, \quad \text{or} \quad 5. \ f \in C,$$

it is necessary and sufficient that 3. $\{S_R(\lambda, A, x)\}$ *converge in the mean with exponent* 1, *or* 4. $|S_R(\lambda, A, x)| \leqq M$, *or* 5. $\{S_R(\lambda, A, x)\}$ *converge uniformly in* B.

We shall write $\mathfrak{S}(\lambda, A, x) \in S$ to indicate that $\mathfrak{S}(\lambda, A, x) = \mathfrak{S}(d\Phi, x)$ with $\int_E |d\Phi| \leqq M$, and $\mathfrak{S}(\lambda, A, x) \in L$, M, or C, if $\mathfrak{S}(\lambda, A, x) = \mathfrak{S}(f, x)$ with $f \in L$, $f \in M$ or $f \in C$ respectively. We will write $\lambda \in (P, Q)$ to indicate that $\mathfrak{S}(I, A, x) \in P$ implies $\mathfrak{S}(\lambda, A, x) \in Q$.

5. **Transformations.** We begin with the following lemma.

LEMMA 1. $\mathfrak{S}(I, I, x) \in S$.

PROOF. Let $f \in C$. Since $S_R(I, I, x) = K_R(x)$, we have

$$S_R(f, x) = B(n) \int_E f(u) S_R(I, I, u - x) du,$$

but f is continuous, the left-hand member converges uniformly to f,

and for any sequence of values of R, $\{|S_R(f, x)|\}$ is bounded uniformly. $\{S_R(f,x)\}$ is a sequence of linear operations whose norms are

$$\int_E |S_R(I, I, x)|\, dx,$$

and by a simple corollary of a theorem of Banach and Steinhaus[4] these must be uniformly bounded, or $\mathfrak{S}(I, I, x) \in S$.

We now have this theorem.

THEOREM 1. *The transformation classes* (S, S), (M, M), (L, L), *and* (C, C) *are identical, and* λ *belongs to each of them if and only if* $\mathfrak{S}(\lambda, I, x) \in S$, *or*

$$\int_E |S_R(\lambda, I, x)|\, dx \leqq M. \tag{10}$$

(A) $\lambda \in (S, S) \rightarrow (10)$. This follows immediately from Lemma 1.

(B) $(10) \rightarrow \lambda \in (S, S)$. From (5) we have

$$\begin{aligned} S_R(\lambda, A, x) &= \sum_{(k)} \Psi\left(\frac{|K|}{R}\right) \lambda(k) A(k) e^{i(k\cdot x)} \\ &= \sum_{(k)} \Psi\left(\frac{|K|}{R}\right) \lambda(k) (2\pi)^{-n} \int_{E(u)} e^{i(k\cdot x - u)} d\Phi(E)^{5} \\ &= (2\pi)^{-n} \int_{E(u)} S_R(\lambda, I, x-u) f(u) d\Phi(E), \end{aligned}$$

whence

$$\int_E |S_R(\lambda, A, x)|\, dx \leqq (2\pi)^{-n} \cdot M \cdot \int_E |d\Phi|.$$

(C) $\lambda \in (M, M) \rightarrow (10)$ and $\lambda \in (C, C) \rightarrow (10)$.

We have immediately in both cases

$$S_R(\lambda, A, x) = B(n) \int_{E(u)} S_R(\lambda, I, x-u) f(u) du$$

and the boundedness of the set $\{S_R(\lambda, A, 0)\}$ implies (10).

(D) $(10) \rightarrow \lambda \in (M, M)$, $(10) \rightarrow \lambda \in (C, C)$, and $(10) \rightarrow \lambda \in (L, L)$.

From (5) and (10) we have

---

[4] S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczne, vol. 1, Warsaw, 1932, p. 80.

[5] The notation $\int_{E(u)} f(x, u) d\Phi(E)$ is used to indicate the integral over E of f regarded as a function of a point (u).

$$S_R(\lambda, A, x) = \sum_{(k)} \Psi\left(\frac{|K|}{R}\right) A(k)(2\pi)^{-n} \int_{E(u)} e^{-i(k\cdot x-u)} d\Phi(E)$$

$$= (2\pi)^{-n} \int_{E(u)} S_R(I, A, x-u) d\Phi(E),$$

and the boundedness, uniform convergence, or mean convergence of $\{S_R(I, A, x)\}$ implies the same for $\{S_R(\lambda, A, x)\}$.

(E) $\lambda \in (L, L) \rightarrow (10)$.

Suppose (10) does not hold. Then there is a sequence $\{R_m\}$ and a sequence of sets $\{G_m\}$, each G being a finite sum of nonoverlapping cubes of E, such that

$$\int_{G_m} S_{R_m}(\lambda, A, x)dx = (2\pi)^{-n} \int_{E(u)} f(u) \left\{ \int_{G_m} S_{R_m}(\lambda, I, x-u)dx \right\} du$$

is not bounded. But this is a sequence of linear functionals defined over L, of norm

$$\max_u \left| \int_{G_m} S_{R_m}(\lambda, I, x-u)dx \right|,$$

and by the theorem of Banach and Steinhaus[4] there is an $f \in L$ such that

$$\left\{ \int_E S_{R_m}(\lambda, A, x)dx \right\}$$

is unbounded, and hence $\mathfrak{S}(\lambda, A, x) \notin S$, which obviously implies $\mathfrak{S}(\lambda, A, x) \notin L$, contrary to hypothesis.

THEOREM 2. *The transformation classes (S, L) and (M, C) are identical, and λ belongs to each of them if and only if $\mathfrak{S}(\lambda, I, x) \in L$, or*

$$\lim_{R, R' \to \infty} \int_E \left| S_R(\lambda, I, x) - S_{R'}(\lambda, I, x) \right| dx = 0. \tag{11}$$

PROOF. (A) $\lambda \in (S, L) \rightarrow (11)$, immediately, from Lemma 1.

(B) $(11) \rightarrow \lambda \in (S, L)$.

Since $S_R(\lambda, A, x) = (2\pi)^{-n} \int_{E(u)} S_R(\lambda, I, x-u) d\Phi(E)$ it follows that (11) implies the convergence in the mean of $\{S_R(\lambda, A, x)\}$.

(C) $(11) \rightarrow \lambda \in (M, C)$.

Since $S_R(\lambda, A, x) = (2\pi)^{-n} \int_{E(u)} S_R(\lambda, I, x-u) f(u) du$, with $|f| \leqq M$ by hypotheses, the uniform convergence of the sequence $\{S_R(\lambda, A, x)\}$ follows immediately from (11).

(D) $\lambda \in (M, C) \rightarrow (11)$.

Writing

$$S_R(\lambda, A, x) = (2\pi)^{-n} \int_{E(u)} S_R(\lambda, I, x - u) f(u) du,$$

we have

$$S_R(\lambda, A, 0) = (2\pi)^{-n} \int_E S_R(\lambda, I, u) f(u) du = U_f(S_R(\lambda, I, u))$$

which defines a linear function U_f. The sequence of values $\{U_f(S_R(\lambda, I, u))\}$ converges for every bounded f, that is, $\{S_R(\lambda, I, u)\}$ converges weakly. But L is weakly complete,[6] whence there is an $F \in L$ such that

$$\lim_{R\to\infty} \int_E \{S_R(\lambda, I, u) - F(u)\} f(u) du = 0$$

for every bounded f. Consider now the two series

$$\mathfrak{S}(F, u) \simeq \sum_{(k)} C(k) e^{i(k\cdot u)}$$

and

$$\mathfrak{S}(\lambda, I, u) \simeq \sum_{(k)} \lambda(k) e^{i(k\cdot u)},$$

and note that

$$S_R(\lambda, I, u) = \sum_{(k)} \Psi\left(\frac{|K|}{R}\right) \lambda(k) e^{i(k\cdot u)}.$$

Let

$$U_N(g) = \int_E g(u) e^{-i(N\cdot u)} du,$$

which is linear on L. Now

$$\lim_{R\to\infty} U_N(S_R(\lambda, I, u)) = \lim_{R\to\infty} \Psi\left(\frac{|N|}{R}\right) \lambda(N) = \lambda(N)$$

$$= U_N(F) = C(N)$$

whence

$$S_R(\lambda, I, u) = S_R(F, u).$$

But

[6] S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczne, vol. 1, Warsaw, 1932, p. 141.

$$\lim_{R\to\infty}\int_E \left| S_R(F, u) - F(u) \right| du = 0$$

which, in view of the preceding equality, implies (11).

THEOREM 3. *The transformation classes* (S, M), (L, C), *and* (L, M) *are identical, and* λ *belongs to each of them if and only if* $\mathfrak{S}(\lambda, I, x)\in M$, *or*

$$|S_R(\lambda, I, x)| \leqq M. \tag{12}$$

PROOF. (A) $\lambda\in(S, M)\rightarrow(12)$ follows immediately from Lemma 1.

(B) $(12)\rightarrow\lambda\in(S, M)$.

Since

$$S_R(\lambda, A, x) = \int_{E(u)} S_R(\lambda, I, x - u)d\Phi(E),$$

we have from (12)

$$|S_R(\lambda, A, x)| \leqq M\int_E |d\Phi(E)|, \quad \text{and} \quad \mathfrak{S}(\lambda, A, x)\in M.$$

(C) $(12)\rightarrow\lambda\in(L, C)$.

We have immediately

$$|S_R(\lambda, A, x) - S_{R'}(\lambda, A, x)| \leqq M\int_E |S_R(I, A, x) - S_{R'}(I, A, x)|dx$$

where $\mathfrak{S}(\lambda, I, x) = \mathfrak{S}(f, x)$, and $|f| \leqq M$. The integral on the right approaches 0 as R and R' approach infinity, whence $\mathfrak{S}(\lambda, A, x)\in C$.

(D) $\lambda\in(L, C)\rightarrow(12)$.

We write immediately

$$(2\pi)^n S_R(\lambda, A, x) = \int_{E(u)} S_R(\lambda, I, x - u)f(u)du$$

and (12) follows from a theorem of Steinhaus and Banach.[4]

(E) $\lambda\in(L, C)\rightarrow\lambda\in(L, M)$ is obvious.

(F) $\lambda\in(L, M)\rightarrow\lambda\in(L, C)$.

Since

$$S_R(\lambda, A, x) = (2\pi)^{-n}\int_{E(u)} S_R(\lambda, I, x - u)f(u)du$$

and $|S_R(\lambda, A, x)| \leqq M(f)$ by hypothesis, it follows that $S_R(\lambda, A, x)$ exists for every $f\in L$, and $\{S_R(\lambda, A, 0)\}$ is a sequence of linear functionals on L, whose norms are

$$\operatorname*{ess\,sup}_{u} \left| S_R(\lambda, I, u) \right| = \max_{u} \left| S_R(\lambda, I, u) \right|,$$

and for every $f \in L$

$$\left| S_R(\lambda, A, 0) \right| \leqq M(f),$$

whence, by the theorem of Steinhaus and Banach[4] the norms are uniformly bounded. Hence $\lambda \in (L, M) \rightarrow (12) \rightarrow \lambda \in (L, C)$.

THEOREM 4. *λ belongs to the transformation class (S, C) if and only if $\mathfrak{S}(\lambda, I, x) \in C$, or*

$$\lim_{R,\, R \to \infty} \left| S_R(\lambda, I, x) - S_{R'}(\lambda, I, x) \right| = 0, \quad \textit{uniformly in } x \in E. \tag{13}$$

PROOF. (A) $\lambda \in (S, C) \rightarrow (13)$ immediately from Lemma 1.

(B) $(13) \rightarrow \lambda \in (S, C)$.

Since

$$S_R(\lambda, A, x) - S_{R'}(\lambda, A, x) = (2\pi)^{-n} \int_{E(u)} S_R(\lambda, I, x - u) d\Phi(E) - (2\pi)^{-n} \int_{E(u)} S_{R'}(\lambda, I, x - u) d\Phi(E),$$

we may write

$$\left| S_R(\lambda, A, x) - S_{R'}(\lambda, A, x) \right| \leqq \sup_{x \in E} \left| S_R(\lambda, I, x) - S_{R'}(\lambda, I, x) \right| \int_E \left| d\Phi(E) \right|$$

and the uniform convergence of $\{S_R(\lambda, A, x)\}$ follows from that of $\{S_R(\lambda, I, x)\}$.

6. **Conclusion.** Of the factor sequence transformations among the classes S, L, M, and C, all of those characterizable in terms of these classes applied to $\mathfrak{S}(\lambda, I, x)$ have been discussed. The class (L, M) of transformations does not exist in a proper sense since, by Theorem 3, its range is a subset of C contained in M. The results of the paper may be taulated as follows:

$$\mathfrak{S}(\lambda, I, x) \in S \equiv \lambda \in (S, S) \equiv (L, L) \equiv (M, M) \equiv (C, C),$$
$$\mathfrak{S}(\lambda, I, x) \in L \equiv \lambda \in (S, L) \equiv (M, C),$$
$$\mathfrak{S}(\lambda, I, x) \in M \equiv \lambda \in (S, M) \equiv (L, C),$$
$$\mathfrak{S}(\lambda, I, x) \in C \equiv \lambda \in (S, C).$$

BROWN UNIVERSITY AND THE CITADEL

SOME PROPERTIES OF MEASURABLE FUNCTIONS

H. FEDERER AND A. P. MORSE

1. **Introduction.** Throughout this paper the letter I will denote some fixed closed interval and f a numerically valued measurable function on I. It is our purpose to establish certain general properties of f. We point out in §4 that two theorems of Banach are almost immediate consequences of these properties. We suspect that further use can be made of our results.

2. **Some notations.** We define

$$X^{\wedge} = \underset{y}{E}\,[y = f(x) \text{ for some } x \in X], \qquad X \subset I,$$

$$Y^{\vee} = \underset{x}{E}\,[f(x) \in Y], \qquad Y \subset I^{\wedge}.$$

Writing $X^{\wedge\vee} = (X^{\wedge})^{\vee}$ and $Y^{\vee\wedge} = (Y^{\vee})^{\wedge}$ we note that the relations

$$X \subset X^{\wedge\vee}, \qquad Y = Y^{\vee\wedge},$$

$$\Big(\sum_{n=1}^{\infty} X_n\Big)^{\wedge} = \sum_{n=1}^{\infty} X_n^{\wedge}, \qquad \Big(\sum_{n=1}^{\infty} Y_n\Big)^{\vee} = \sum_{n=1}^{\infty} Y_n^{\vee},$$

$$\Big(\prod_{n=1}^{\infty} X_n\Big)^{\wedge} \subset \prod_{n=1}^{\infty} X_n^{\wedge}, \qquad \Big(\prod_{n=1}^{\infty} Y_n\Big)^{\vee} = \prod_{n=1}^{\infty} Y_n^{\vee},$$

$$X_1^{\wedge} - X_2^{\wedge} \subset (X_1 - X_2)^{\wedge}, \qquad Y_1^{\vee} - Y_2^{\vee} = (Y_1 - Y_2)^{\vee},$$

$$YX^{\wedge} = (Y^{\vee}X)^{\wedge},$$

hold whenever $X, X_1, X_2, \cdots$ are subsets of I and $Y, Y_1, Y_2, \cdots$ are subset of $I^{\wedge}$.

We further define

$$\{y\} = \underset{z}{E}\,[z = y],$$

$$\mathfrak{F} = \underset{y}{E}\,[\{y\}^{\vee} \text{ is finite}],$$

$$\mathfrak{K} = \underset{y}{E}\,[\{y\}^{\vee} \text{ has at least } \aleph_0 \text{ elements}],$$

$$\mathfrak{Q} = \underset{y}{E}\,[\{y\}^{\vee} \text{ has more than } \aleph_0 \text{ elements}].$$

Received by the editors June 4, 1942.

We denote the (outer) linear Lebesgue measure of a set A by $|A|$.

3. **Some theorems concerning f.**

THEOREM 3.1. *If A is a measurable subset of I, then there is a measurable subset B of A such that $B^{\wedge}=A^{\wedge}$ and f is univalent on B.*

PROOF. Let[1] $0=C_0\subset C_1\subset C_2\subset\cdots$ be closed subsets of A relative to each of which f is continuous and for which

$$C=\sum_{n=1}^{\infty}C_n, \qquad |A-C|=0.$$

For each positive integer n let[2] D_n be a measurable subset of C_n such that $D_n^{\wedge}=C_n^{\wedge}$ and f is univalent on D_n. Let

$$B_1=\sum_{n=1}^{\infty}[D_n-C_{n-1}^{\wedge\vee}].$$

Noting that $C_{n-1}^{\wedge}$ and $C_nC_{n-1}^{\wedge\vee}$ are closed and also that

$$D_n-C_{n-1}^{\wedge\vee}=D_n-C_nC_{n-1}^{\wedge\vee}, \qquad n=1,2,3,\cdots,$$

we see that B_1 is a measurable subset of C.

Now let $y_0\in C^{\wedge}$. Let k be the least integer in

$$\underset{n}{E}\,[y_0\in C_n^{\wedge}].$$

Since

$$y_0\in C_k^{\wedge}-C_{k-1}^{\wedge}=D_k^{\wedge}-C_{k-1}^{\wedge}\subset(D_k-C_{k-1}^{\wedge\vee})^{\wedge},$$

take a number x_0 such that $f(x_0)=y_0$ and

$$x_0\in[D_k-C_{k-1}^{\wedge\vee}]\subset B_1.$$

Thus $y_0\in B_1^{\wedge}$. Moreover if $x_0\neq x\in B_1$ with $f(x)=y_0$, we could, for some integer $l>k$, successively infer the relations

$$x\in D_l-C_{l-1}^{\wedge\vee}, \qquad y_0\in C_k^{\wedge}\subset C_{l-1}^{\wedge}. \qquad x\in C_{l-1}^{\wedge\vee},$$

the first and last of which contradict.

Consequently $C^{\wedge}\subset B_1^{\wedge}$, f is univalent on B_1, $C^{\wedge}=B_1^{\wedge}$.

Let $\alpha=A-C^{\wedge\vee}$ and select a set $B_2\subset\alpha$ such that $B_2^{\wedge}=\alpha^{\wedge}$ and f is univalent on B_2. Now

[1] See Saks, *Theory of the integral*, Warsaw, 1937, p. 72.

[2] See Saks, loc. cit. p. 282.

$$B_2 \subset \alpha \subset A - C, \qquad |B_2| = 0.$$

Defining $B = B_1 + B_2$ we see that B is measurable and that

$$A^\wedge \supset B^\wedge = B_1^\wedge + B_2^\wedge$$
$$= C^\wedge + (A - C^{\wedge\vee})^\wedge \supset C^\wedge + (A^\wedge - C^\wedge) = A^\wedge.$$

Lastly f is univalent on B; for otherwise, in view of its univalence on B_1 and B_2, we could select points $x_1 \in B_1$ and $x_2 \in B_2$ with $f(x_1) = f(x_2)$ and deduce the false proposition

$$x_2 \in B_2 B_1^{\wedge\vee} \subset \alpha C^{\wedge\vee} = (A - C^{\wedge\vee}) C^{\wedge\vee} = 0.$$

The proof is complete.

LEMMA 3.2. *If f is continuous relative to each of the closed sets $A_1 \subset A_2 \subset A_3 \subset \cdots$ with*

$$A = \sum_{j=1}^{\infty} A_j \subset I,$$

then

$$S_n = A \cdot \left(\underset{y}{E} \left[A\{y\}^\vee \text{ has at least } n \text{ elements} \right] \right)^\vee$$

is a Borel set for each positive integer n.

PROOF. Let n be a positive integer.

For each positive integer j let W_j be that subset of Euclidean n-space such that $P = (P_1, P_2, \cdots, P_n)$ is in W_j if and only if

$$P_i \in A_j, \qquad i = 1, 2, \cdots, n,$$
$$P_i - P_{i-1} \geqq 1/j, \qquad i = 2, 3, \cdots, n.$$

For $j = 1, 2, \cdots 3, \cdots$ note that W_j is bounded and closed and let

$$B_j = A_j \underset{x}{E} [\text{there is a } P \in W_j \text{ with } f(x) = f(P_i) \text{ for } i = 1, 2, \cdots, n].$$

It follows that

$$S_n = \sum_{j=1}^{\infty} B_j. \tag{1}$$

Now let j_0 be a positive integer and $x_1, x_2, x_3, \cdots$ be points of B_{j_0} such that

$$\lim_{m\to\infty} x_m = x_0 \in I.$$

Clearly $x_0 \in A_{j_0}$. There are points $P^1, P^2, P^3, \cdots$ of W_{j_0} such that

$$f(x_m) = f(P_i^m), \qquad m = 1, 2, 3, \cdots; \; i = 1, 2, 3, \cdots, n.$$

Let α be a set of integers and P^0 be a point of the compact set W_{j_0} such that

$$\lim_{m\to\infty,\, m\in\alpha} P^m = P^0.$$

Since f is continuous relative to A_{j_0} we conclude

$$f(P_i^0) = \lim_{m\to\infty,\, m\in\alpha} f(P_i^m) = \lim_{m\to\infty,\, m\in\alpha} f(x_m) = f(x_0), \qquad i = 1, 2, \cdots, n,$$

which implies $x_0 \in B_{j_0}$. Hence B_{j_0} is closed.

In view of (1) the proof is complete.

THEOREM 3.3. *There is a measurable*[3] *set C such that $C^\wedge = \mathfrak{R}$ and*

$$(\{y\}^\vee - C) \text{ is finite for each } y \in I^\wedge.$$

PROOF. Retaining the notation of the statement of 3.2 and demanding in addition that $|I - A| = 0$, we define

$$S = A \cdot \left(\underset{y}{E} \left[A\{y\}^\vee \text{ is infinite} \right] \right)^\vee, \tag{2}$$

$$T = (I - A) \cdot \left(\underset{y}{E} \left[(I - A)\{y\}^\vee \text{ is infinite} \right] \right)^\vee, \tag{3}$$

$$C = S + T.$$

The set C is measurable because $S = \prod_{n=1}^{\infty} S_n$ and $|T| = 0$. The fact that $C^\wedge \subset \mathfrak{R}$ may be easily verified by deleting A and $(I - A)$ from (2) and (3), respectively.

The assumption that $(A\{y\}^\vee - C)$ is infinite leads to the relations

$$A\{y\}^\vee \subset S \subset C, \qquad A\{y\}^\vee - C = 0,$$

the second of which is contradictory. Similarly the assumption that $((I - A)\{y\}^\vee - C)$ is infinite leads to a contradiction.

Thus $\{y\}^\vee - C$ is finite for each $y \in I^\wedge$, $C\{y\}^\vee$ is infinite (and non-

[3] Using the fact that there is a perfect set of measure zero which is decomposable into a continuum of disjoint perfect sets, it can be shown that a judicious choice of our function f insures that neither $\mathfrak{R}^\vee$ nor $\mathfrak{Q}^\vee$ is measurable.

vacuous) for $y \in \mathfrak{R}$; consequently $\mathfrak{R} \subset C^\wedge$.

Hence $\mathfrak{R} = C^\wedge$.

THEOREM 3.4.[4] *Let $\epsilon > 0$. Then*

(i) *there is a measurable set $L \subset I$ with $L^\wedge = \mathfrak{R}$ and $|L| < \epsilon$;*

(ii) *there is a measurable set $R \subset I$ with $R^\wedge = \mathfrak{Q}$ and $|R| = 0$.*

PROOF. Let C be a measurable set such that $C^\wedge = \mathfrak{R}$ and $\{y\}^\vee - C$ is finite for each $y \in \mathfrak{R}$.

Let F be the family of measurable sets $X \subset C$ such that $X\{y\}^\vee$ is finite for each $y \in X^\wedge$. Note that $0 \in F \neq 0$ and define

$$M = \sup_{X \in F} |X|$$

and let $X_1, X_2, X_3, \cdots$ be sets in F such that $\lim_{n\to\infty} |X_n| = M$. Letting

$$S_n = \sum_{j=1}^{n} X_j, \qquad n = 1, 2, 3, \cdots,$$

observe that $X_n \subset S_n \in F$. Thus $|X_n| \leqq |S_n| \leqq M$ for $n = 1, 2, 3, \cdots$ and consequently

$$|S| = \lim_{n\to\infty} |S_n| = M \qquad \text{where } S = \sum_{n=1}^{\infty} S_n.$$

Let n_0 be an integer such that $|S - S_{n_0}| < \epsilon$. Let T be a measurable (3.1) subset of $C - S$ such that $T^\wedge = (C - S)^\wedge$ and f is univalent on T.

Defining

$$L = T + (S - S_{n_0}), \quad R = T\mathfrak{Q}^\vee,$$

we verify first that $|T| = 0$. If $|T|$ were greater than 0 there would be an integer k such that

$$|S_k| > M - |T|, \qquad (T + S_k) \in F, \qquad TS_k \subset TS = 0,$$

$$M \geqq |T + S_k| = |T| + |S_k| > M, \qquad M > M.$$

Consequently L and R are measurable, $|L| < \epsilon$, $|R| = 0$.

Next since $S_{n_0}\{y\}^\vee$ is at most finite, $S\{y\}^\vee$ is at most countable and $C\{y\}^\vee$ has the same power as $\{y\}^\vee$ for $y \in \mathfrak{R}$, we see that

$$\mathfrak{R} \subset (C - S_{n_0})^\wedge \subset \mathfrak{R}, \qquad \mathfrak{Q} \subset (C - S)^\wedge = T^\wedge.$$

From these relations it follows almost immediately that $L^\wedge = \mathfrak{R}$ and $R^\wedge = \mathfrak{Q}$.

[4] It is evident from Theorem 3.1 that Theorem 3.4 remains true if we add the requirement that f be univalent on L and on R.

4. **Some applications.** We say our function f satisfies condition S if and only if for each $\epsilon>0$ there is a $\delta>0$ such that

$$|X^\wedge\cdot[-1/\epsilon, 1/\epsilon]|<\epsilon \quad \text{whenever } X\subset I,\ |X|<\delta;$$

f satisfies condition N if and only if

$$|X^\wedge|=0 \quad \text{whenever } X\subset I,\ |X|=0;$$

f satisfies condition T_1 if and only if $|\mathfrak{R}|=0$; and f satisfies condition T_2 if and only if $|\mathfrak{Q}|=0$.

The following two theorems of Banach[5] are essentially corollaries of 3.4.

THEOREM 4.1. *If f satisfies condition N, then it satisfies T_2.*

THEOREM 4.2. *A necessary and sufficient condition that f satisfy condition S is that it satisfy both N and T_1.*

Theorem 4.1 and the *necessity* in 4.2 are immediate consequences of 3.4. The *sufficiency* in 4.2 may be proved as follows:

Suppose f satisfies N and T_1 but not S. Then we can find a number $\epsilon>0$ and measurable subsets $A_1\supset A_2\supset A_3\supset\cdots$ of I such that

$$\lim_{n\to\infty}|A_n|=0, \qquad |A_n^\wedge|\geqq\epsilon, \qquad A_n^\wedge\subset[-1/\epsilon, 1/\epsilon],$$
$$n=1, 2, 3, \cdots.$$

We note[6] that $A_n^\wedge$, $I^\wedge$, $\mathfrak{R}$, $\mathfrak{F}$, are measurable and define

$$A=\prod_{n=1}^{\infty}A_n.$$

Since the product of a descending sequence of nonvacuous *finite* sets is nonvacuous, we have

$$\mathfrak{F}\prod_{n=1}^{\infty}A_n^\wedge\subset\left(\prod_{n=1}^{\infty}A_n\right)^\wedge=A^\wedge.$$

Hence

$$|A^\wedge|\geqq\left|\mathfrak{F}\prod_{n=1}^{\infty}A_n^\wedge\right|=\lim_{n\to\infty}|\mathfrak{F}A_n^\wedge|=\lim_{n\to\infty}|A_n^\wedge|\geqq\epsilon$$

in contradiction to the relations

$$|A|=\lim_{n\to\infty}|A_n|=0, \qquad |A^\wedge|=0.$$

---

[5] See Saks, loc. cit. p. 284.

[6] If f satisfies N, then $X^\wedge$ is measurable, whenever X is a measurable subset of I.

Theorem 4.2 is proved.

5. **Generalizations.** Suppose E is a metric space and ϕ is a measure over[7] E such that: closed sets are ϕ measurable in the sense of Carathéodory;[8] every ϕ measurable set is the sum of an F_σ and a set of ϕ measure zero; E is a countable sum of compact sets; $\phi(E) < \infty$.

Let H be another metric space. We say a function q is ϕ finitely valued, if it is on E, its range is a finite subset of H and

$$\underset{x}{E}\,[q(x) = y]$$

is ϕ measurable for each $y \in H$. We call a function ϕ measurable, if it is ϕ almost everywhere in E the limit of a sequence of ϕ finitely valued functions.

To generalize our results we replace I by E, Lebesgue measure by ϕ, f by a ϕ measurable function g, the word "closed" in the statement of 3.2 by "compact." Conditions S, N, T_1, T_2 are to be interpreted in terms of such a measure ψ over H that closed subsets of H are ψ measurable and bounded sets have finite ψ measure.

All our theorems remain true under these conditions with properly adjusted notation. Leaving details to the reader, we solve the only non-trivial problem arising in this extension by the following:

THEOREM 5.1. *If g is continuous relative to the compact set $C \subset E$, then there is a Borel set $B \subset C$ such that $B^\wedge = C^\wedge$ and g is univalent on B.*

PROOF. Select[9] a continuous function h whose domain is a perfect set $A \subset [0, 1]$ and whose range is C.

For $n = 1, 2, 3, \cdots$, let A_n be the set such that $t \in A_n$ if and only if $t \in A$ and the relations

$$s \in A, \quad s \leqq t - n^{-1}, \quad g[h(s)] = g[h(t)],$$

are incompatible. Note that A_n is open with respect to A. Let

$$B_n = C \cdot \underset{x}{E}\,[x = h(t) \text{ for some } t \in A_n],$$

$$B = \prod_{n=1}^{\infty} B_n.$$

Observe that B is an $F_{\sigma\delta}$.

---

[7] We say Λ is over E if and only if Λ is a function whose domain is the set of subsets of E.

[8] See H. Hahn, *Theorie der reellen Funktionen*, vol. 1, Berlin, 1921, p. 424.

[9] See W. Sierpinski, *Introduction to general topology*, Toronto, 1934, p. 166.

Now choose $y_0 \in C^{\wedge}$. Taking

$$t_0 = \inf A \cdot \underset{t}{E} [g[h(t)] = y_0], \qquad x_0 = h(t_0),$$

we easily check that

$$t_0 \in \prod_{n=1}^{\infty} A_n, \qquad x_0 \in B, \qquad g(x_0) = y_0.$$

If $x_0 \neq x_1 \in B$ with $g(x_1) = y_0$, we infer that

$$T = A \cdot \underset{t}{E} [h(t) = x_1]$$

is a closed set lying entirely to the right of t_0 with distance $d > 0$ from t_0; hence $n > d^{-1}$ implies $T \subset A - A_n$ which in turn implies $x_1 \in E - B_n \subset E - B$ contrary to the assumption $x_1 \in B$.

Theorem 5.1 is proved.

Now we replace the condition that $\phi(E) < \infty$ by the hypothesis:

$$E = \sum_{n=1}^{\infty} E_n, \qquad E_n E_m = 0 \quad \text{for} \quad n \neq m,$$

E_n is ϕ measurable with $\phi(E_n) < \infty$.

Obviously Theorem 3.4 part (i) and Theorem 4.2 do not[10] hold under these new conditions. However 3.1, 3.2, 3.3, 3.4 part (ii), 4.1 are still valid. To prove this it is sufficient to know of the existence of a measure Φ over E such that: $\Phi(E) < \infty$; a set is Φ measurable if and only if it is ϕ measurable; a set has Φ measure zero if and only if it has ϕ measure zero.

Let $\lambda_1, \lambda_2, \lambda_3, \cdots$ be *positive* numbers such that

$$\sum_{n=1}^{\infty} \lambda_n \phi(E_n) < \infty$$

and define Φ over E by the relation:

$$\Phi(X) = \sum_{n=1}^{\infty} \lambda_n \phi(X E_n) \qquad \text{for } X \subset E.$$

The reader will find no difficulty in checking that Φ is a measure with the required properties.

UNIVERSITY OF CALIFORNIA

[10] Counterexample: $g(x) = \sin x$, $-\infty < x < +\infty$.

SOME THEOREMS ON THE EULER ϕ-FUNCTION

N. G. GUNDERSON

The Euler ϕ-function, $\phi(m)$, denotes the number of positive integers not greater than m which are relatively prime to m.[1] It was noted by U. Scarpis[2] that $n \mid \phi(p^n-1)$. Generalizations of this result are obtained in Theorems 9 and 10.

The first five theorems are either well known or self-evident.[3]

THEOREM 1. *If $p_1, \cdots, p_k$ are the distinct prime factors of m, then*

$$\phi(m) = m(p_1 - 1)(p_2 - 1) \cdots (p_k - 1)/p_1 p_2 \cdots p_k.$$

THEOREM 2. *If $a_1, \cdots, a_k$ are relatively prime in pairs, then*

$$\phi(a_1 \cdots a_k) = \phi(a_1) \cdot \phi(a_2) \cdots \phi(a_k).$$

THEOREM 3. *If w is the product of the distinct prime factors common to m and n, then*

$$\phi(mn) = w \cdot \phi(m) \cdot \phi(n)/\phi(w).$$

THEOREM 4. *If $a \mid b$, then $\phi(a) \mid \phi(b)$.*

THEOREM 5. *If $q \mid a$ and $q \equiv 1 \pmod{p^\alpha}$, then $p^\alpha \mid \phi(a)$.*

THEOREM 6. *If p is an odd prime, $a \neq b$, and $\alpha \geqq 1$, then*

$$p^{2\alpha-1} \mid \phi(a^{p^\alpha} + b^{p^\alpha}).$$

The proof is by induction on α. We assume $a > b$. In the notation of Birkhoff and Vandiver,[4] $a^p + b^p = V_{2p}/V_p$. By their Theorems V and I, there is a prime divisor q of $a^p + b^p$ such that $q \equiv 1 \pmod{p}$ unless $p = 3$, $a = 2$, $b = 1$. Then by Theorem 5, $p \mid \phi(a^p + b^p)$, and in the exceptional case, $3 \mid \phi(2^3 + 1^3)$. Thus the theorem holds for $\alpha = 1$, starting the induction, so we assume it for all positive integers less than α. We adopt the notation $C = AB$, where

$$C = a^{p^\alpha} + b^{p^\alpha}, \qquad P = p^{\alpha-1}, \qquad A = a^p + b^p,$$

$$B = a^{(p-1)P} - a^{(p-2)P} \cdot b^P + \cdots - a^P \cdot b^{(p-2)P} + b^{(p-1)P}.$$

Received by the editors June 12, 1942.

[1] In this discussion all letters represent positive integers. In particular, p and q represent primes.

[2] Period. Mat. vol. 29 (1913) p. 138.

[3] See, for example, L. E. Dickson, *History of the theory of numbers*, vol. 1, chap. 5.

[4] Ann. of Math. (2) vol. 5 (1903) pp. 173–180.

Case 1. $(a, b)=1$.

Again using the notation of Birkhoff and Vandiver, $B=C/A=(V_{2pP}/V_{pP})(V_P/V_{2P})$, so we see by their Theorems V and I that there is a prime divisor q of B such that $q\equiv 1 \pmod{p^\alpha}$. Hence by Theorem 5 we have that $p^\alpha \mid \phi(B)$. By the hypothesis of induction we have $p^{2(\alpha-1)-1} \mid \phi(A)$.

Now, if $(A, B)\neq 1$, let r be a common prime factor of A and B. Then $a^P \equiv -b^P \pmod r$, so that $B\equiv p\cdot a^{(p-1)P}\equiv 0 \pmod r$. If $r \mid a$, then $r \mid b$, contrary to $(a, b)=1$, so $r=p$. Then by Theorem 3,

$$\phi(C) = p\cdot\phi(A)\cdot\phi(B)/(p-1),$$

so we have that $p\cdot p^{2(\alpha-1)-1}\cdot p^\alpha \mid \phi(C)$. But since $\alpha>1$, $3\alpha-2>2\alpha-1$, so $p^{3\alpha-1} \mid \phi(C)$.

If $(A, B)=1$, then $\phi(C)=\phi(A)\cdot\phi(B)$, and $p^{2(\alpha-1)-1}\cdot p^\alpha \mid \phi(C)$. Again, since $\alpha>1$, $3\alpha-3\geqq 2\alpha-1$, and $p^{3\alpha-1} \mid \phi(C)$.

Case 2. $(a, b)\neq 1$.

Let $(a, b)=c$, $a=ca_1$, $b=cb_1$, $(a_1, b_1)=1$. Further, since $a>b$, a_1 and b_1 are not both 1, and so $a_1>b_1$. By Case 1, $p^{2\alpha-1} \mid \phi(a_1^{p^\alpha}+b_1^{p^\alpha})$ so by Theorem 4, $p^{2\alpha-1} \mid \phi(a^{p^\alpha}+b^{p^\alpha})$.

THEOREM 7. *If* $a\neq b$, *then* $2^{\alpha+1} \mid \phi(a^{2^\alpha}+b^{2^\alpha})$.

We note that $a^{2^\alpha}+b^{2^\alpha}=2^\beta$ would imply $a=b=2$, so $a^{2^\alpha}+b^{2^\alpha}$ has an odd factor, say q. For $(a, b)=1$, Euler[5] has shown that $q\equiv 1 \pmod{2^{\alpha+1}}$, so by Theorem 5, $2^{\alpha+1} \mid \phi(a^{2^\alpha}+b^{2^\alpha})$. For $(a, b)\neq 1$ we proceed as in Case 2 of Theorem 6.

THEOREM 8. *If* $a>b$, *then* $p^{2\alpha-1} \mid \phi(a^{p^\alpha}-b^{p^\alpha})$.

The proof of this theorem parallels that of Theorem 6.

THEOREM 9. *If* $a>b$, *and* m *is the product of the distinct prime factors of* n, *then* $(n^2/m) \mid \phi(a^n-b^n)$.

Let $n=P_1P_2\cdots P_k$ where $P_i=p_i^{\alpha_i}$. Then $a^n-b^n=(a^{n_i})^{P_i}-(b^{n_i})^{P_i}$ where $n_i=P_1\cdots P_{i-1}P_{i+1}\cdots P_k$. Then by Theorem 8, $p_i^{2\alpha_i-1} \mid \phi(a^n-b^n)$, and the theorem follows immediately.

THEOREM 10. *If* $a\neq b$, *then* $n \mid \phi(a^n+b^n)$.

This proof parallels that of Theorem 9.

Theorems 9 and 10 can be combined in various ways, for example, we have this theorem:

---

[5] Commentationes Arithmeticae Collectae, vol. 1, p. 55. See also Amer. Math. Monthly vol. 10 (1903) p. 171, misprint of $2m$ for 2^m.

THEOREM 11. *If* $(a, b) = 1$, *and* m *is the product of the distinct prime factors of* n, *then*

$$(n^2/m) \mid \phi(a^{2n} - b^{2n}).$$

CORNELL UNIVERSITY

APPLICATIONS OF TRANSITIVITIES OF BETWEENNESS IN LATTICE THEORY[1]

M. F. SMILEY AND W. R. TRANSUE

Introduction. This paper solves three characterization problems for lattices[2] [1]. Problem I is to characterize those metric spaces [2] into which lattice operations which are consistent with the given metric [1, p. 41] may be introduced. Problem II is to characterize those members of a rather general class of abstract systems which are modular lattices, while Problem III consists in the characterization of lattices in an even larger class of abstract systems.

Problem I has already been solved by V. Glivenko [3]. He showed that the property: "Among those elements metrically between [4, p. 76; 2] two elements a and b, the element $a \cup b$ is farthest from O," and its dual characterize those metric spaces which are also metric lattices with the same metric and least element O. Our approach to Problem I is through the existence of certain metric singularities [2, p. 47] in every metric lattice. Our solution also involves certain five point transitivities [5, Part I] of metric betweenness. The abstract system involved in Problem II (Problem III) is a wide generalization of the concept of a metric space—so general, in fact, that it also includes the concept of a modular lattice (lattice). We find it not difficult to extend the ideas essential to our solution of Problem I to give analogous solutions of Problems[3] II and III. Briefly, our results consist in characterizing the three important systems: metric

---

Presented to the Society, April 4, 1942; received by the editors June 2, 1942.

[1] This paper is an expansion of our note *Metric lattices as singular metric spaces* which was presented to the Society on December 29, 1941, and which was to appear in Bull. Amer. Math. Soc.

[2] Numbers in square brackets refer to the list of references at the end of the paper.

[3] It would be interesting to characterize metric spaces among our general systems. This problem appears difficult to us, and we make no attempt to solve it.

lattices, modular lattices, and lattices within three increasingly general abstract systems.

1. **A characterization of metric lattices.** In this section we shall give our solution of Problem I. The metric singularities which we shall encounter are *pseudo-linear quadruples* [2, p. 48]. Such a quadruple is one which cannot be imbedded in the Euclidean line, but is such that each of the four triples chosen from it can be so imbedded. In metric lattices such configurations abound. Every pair of elements which are not comparable together with their meet and join form such a quadruple. But the existence of "sufficiently many" pseudo-linear quadruples in a metric space (M, δ) is not enough to ensure that lattice operations can be introduced in such a way that (M, δ) becomes a metric lattice $(M, \delta, <)$. However, the additional assumption of a weak form of either of two five point transitivities for the metric betweenness of (M, δ) suffices.

Before we proceed to prove these statements, let us agree on the following matters of notation. The symbols (M, δ) represent a metric space M with metric δ. For elements $a, b, c \in M$ we use the notation abc to indicate that b is metrically between a and c [4, p. 75; 5, Part II], that is, that $\delta(a, c) = \delta(a, b) + \delta(b, c)$. This relation has the fundamental properties:[4]

$$(\alpha) \qquad abc \leftrightarrow cba.$$

$$(\beta) \qquad abc \cdot acb \leftrightarrow b = c.$$

We shall frequently refer to the transitivities:

$$(t_1) \qquad abc \cdot adb \rightarrow dbc.$$

$$(t_2) \qquad abc \cdot adb \rightarrow adc.$$

which are valid for the metric betweenness of M [4, p. 76; 5, Part II]. If (M, δ) is also a metric lattice $(M, \delta, <)$, then the relation abc is equivalent [3] to the equations

$$(1.1) \qquad (a \cap b) \cup (b \cap c) = b = (a \cup b) \cap (b \cup c).$$

These equations define, in fact, a relation of betweenness in arbitrary lattices [5, Part II]. We shall also use the notation abc to indicate that the equations (1.1) are valid, even when no metric enters the

[4] The second of these statements is intended to be read as, "For each three elements $a, b, c \in M$, the relations abc and acb hold if and only if $b = c$." This convenient notation was used with profit in [5]. We shall attempt to preserve as many of the other notations introduced in [5] as possible.

discussion. In this paper we shall need the fact that this relation of lattice betweenness satisfies (α), (β), (t_1), and

$$(1.2) \qquad abc \rightarrow a \cap c \leqq b \leqq a \cup c.$$

Further properties of this relation are given in [5, Part II].

With reference to each point $p \in M$ of a metric space (M, δ) we may define a partial ordering of M as follows [3].

If $a, b \in M$, then $a <_p b$ in case pab and $a \neq b$. The transitivity of the relation $<_p$ follows from the transitivity t_2 of metric betweenness. The antisymmetry of $\leqq_p$ is apparent. We shall be interested in the following conditions [5, Part I].

$$(T_5^p) \qquad pbc \cdot pdc \cdot bxd \rightarrow pxc.$$

$$(T_7^p) \qquad pbc \cdot pbd \cdot cxd \rightarrow pbx.$$

REMARK 1. Condition (T_5^p) states that an upper bound for the set $[b, d]$ is greater than every element between b and d, while condition (T_7^p) states that a lower bound for the set $[c, d]$ is smaller than every element between c and d.

We shall also need the following condition.

(A^p) *If $a, b \in M$ and if a and b are not comparable by the relation $\leqq_p$, then there are elements $a \cup b$, $a \cap b \in M$ for which we have $(a \cup b)a(a \cap b)$, $(a \cup b)b(a \cap b)$, $a(a \cup b)b$, $a(a \cap b)b$, and $p(a \cap b)(a \cup b)$.*

We may now state the following theorem.

THEOREM 1. *Let (M, δ) be a metric space. If $O \in M$, then $(M, \delta, <_0)$ is a metric lattice with least element O if and only if (M, δ) satisfies (A^O) and (T_5^O) or (A^O) and (T_7^O).*

PROOF. The necessity of the condition (A^O) is obvious. But both (T_5^p) and (T_7^p) also hold in every lattice if p is a distributive element. We prove this in the following lemma which is a generalization of Theorem 9.4 of [5].

LEMMA 1. *If L is a lattice and p is a distributive element of L, then the conditions (T_5^p) and (T_7^p) are true.*

PROOF. We prove (T_5^p) first. Let $p \in L$ be a distributive element of a lattice L, and consider elements $b, c, d, x \in L$ satisfying pbc, pdc, and bxd. From (1.2) we obtain $b \leqq p \cup c$, $d \leqq p \cup c$, and $x \leqq b \cup d$. Hence $x \leqq p \cup c$. Dually, $x \geqq p \cap c$. Since p is distributive, it follows that $(p \cap x) \cup (x \cap c) = x \cap (p \cup c) = x$, and dually. Thus (T_5^p) holds. To prove (T_7^p), let $p \in L$ be a distributive element and consider elements $b, c, d, x \in L$ satisfying pbc, pbd, and cxd. As before, we have $b \leqq p \cup c$,

$b \leqq p \cup d$, and $x \geqq c \cap d$. Consequently, since p is distributive, $p \cup x \geqq p \cup (c \cap d) = (p \cup c) \cap (p \cup d) \geqq b$. Dually, $p \cap x \leqq b$. Another application of the distributivity of p yields $(p \cap b) \cup (b \cap x) = b \cap (p \cup x) = b$. By duality we obtain the relation pbx. Thus (T_7') also holds and the proof of the lemma is complete.

The necessity of the conditions (T_6^0) and (T_7^0) now follows from Lemma 1 and the fact that the least element of a lattice is a distributive element.

We pass to the proof of the sufficiency of our conditions. Using only (A^0) we see that $a \cup b$ and $a \cap b$ are upper and lower bounds, respectively, for the set $[a, b]$. This follows on applying the transitivities t_1 and t_2 to the relations $(a \cup b)a(a \cap b)$, $(a \cup b)(a \cap b)O$, and to the relations $(a \cup b)b(a \cap b)$, $(a \cup b)(a \cap b)O$. The force of the conditions (T_6^0) and (T_7^0) is that each ensures that $a \cup b$ and $a \cap b$ are *unique* least upper and greatest lower bounds, respectively. We shall give the details of the proof only in the case in which (A^0) and (T_6^0) are assumed. The other set of conditions may be handled by almost identical arguments. Hence suppose that the conditions (A^0) and (T_6^0) are valid. We prove first that $a \cup b$ is the least upper bound of the set $[a, b]$ if a and b are not comparable. For, if $z \geqq a, b$ by definition we have Oaz and Obz. Condition (A^0) yields $a(a \cup b)b$, and condition (T_6^0) then gives $O(a \cup b)z$, that is, $z \geqq a \cup b$. The treatment of $a \cap b$ is less trivial. Consider an element $w \in M$ for which $w \leqq a, b$; that is, for which Owa and Owb. We show first that $w > a \cap b$ is impossible as follows. If it were true we should have $O(a \cap b)w$, and this with Owa and Owb would yield, via t_1, the relations $(a \cap b)wa$ and $(a \cap b)wb$. Combining $aw(a \cap b)$ with $a(a \cap b)b$ we would obtain, again by t_1, the relation $w(a \cap b)b$. From this and the relation $(a \cap b)wb$ we would get $(a \cap b)w(a \cap b)$. But this last relation implies that $w = a \cap b$, which is contrary to our assumption. Thus the relation $w > a \cap b$ cannot hold. To show that $w \leqq a \cap b$ we again use indirect proof. If this relation fails to hold, then, since $w \not> a \cap b$, the elements w and $a \cap b$ would not be comparable. The condition (A^0) then implies the existence of elements $w \cup (a \cap b)$ and $w \cap (a \cap b)$ satisfying with w and $a \cap b$ the relations listed in (A^0). The application of the condition (T_6^0) to the relations $O(a \cap b)a$, Owa, and $(a \cap b)((a \cap b) \cup w)w$ gives the relation $O((a \cap b) \cup w)a$. Likewise we obtain $O((a \cap b) \cup w)b$. Applying the transitivity t_1 to $((a \cap b) \cup w)(a \cap b)O$ and $a((a \cap b) \cup w)O$ yields $a((a \cap b) \cup w)(a \cap b)$. Similarly, $b((a \cap b) \cup w)(a \cap b)$. Again, the relations $b((a \cap b) \cup w)(a \cap b)$ and $b(a \cap b)a$ would give, by t_1, the relation $a(a \cap b)((a \cap b) \cup w)$. A final application of the transitivity t_1 to the relations $a(a \cap b)((a \cap b) \cup w)$ and $a((a \cap b) \cup w)(a \cap b)$ would pro-

duce the relation $(a\cap b)((a\cap b)\cup w)(a\cap b)$. But this implies that $a\cap b=(a\cap b)\cup w$, contrary to our assumption that $w\leqq a\cap b$ fails to hold. We conclude that $w\leqq a\cap b$ and that $a\cap b$ is the unique greatest lower bound for the set $[a, b]$. We have now shown that the relation $<_O$ has the lattice property. It is an easy consequence of the condition (A^O) that $\delta(O, a)$ is a sharply positive modular functional and that

$$\delta(a, b) = \delta(a \cup b, O) - \delta(a \cap b, O).$$

This completes the proof that $(M, \delta, <_O)$ is a metric lattice.

REMARK 2. To show that one of the conditions (T_5^O) or (T_7^O) is actually needed in Theorem 1, we exhibit a metric space (M, δ) which satisfies the condition (A^O) but for which the partial orderings $<_q$ fail to have the lattice property for every $q\in M$.

Let M consist of the six points O, a, b, c, d, I, with the distances of distinct points (we omit δ) as follows: $Ia=Ib=Oc=Od=ab=cd=2$, $IO=3$, all other distances between distinct points equal to one. The condition (A^O) is valid for (M, δ) because only the pairs a, b and c, d are not comparable by the relation $<_O$ and for these pairs the choices $c\cup d=I$, $c\cap d=a$ or b, $a\cap b=O$, and $a\cup b=c$ or d are effective in the condition (A^O). Thus $<_O$ is not a lattice ordering. By the symmetry of M it will suffice to show that $<_a$ is not a lattice ordering to prove our assertion completely. In $(M, \delta, <_a)$ we have, besides the obvious fact that a is the least element, only the relations $O<b$, $c<b$, $d<b$, $c<I$, and $d<I$. This is obviously not a lattice ordering of M.

2. **A characterization of modular lattices.** We turn now to a consideration of Problem II. The abstract systems which we shall consider consist of a class K of elements a, b, c, $\cdots$ together with a triadic relation R defined on K and satisfying (α), (β), (t_1), and (t_2). We shall find conditions which are necessary and sufficient that such a system be a modular lattice whose lattice betweenness is identical with the relation R.

For an arbitrary triadic relation R defined on a class K of elements a, b, c, $\cdots$ we make the following definition.

(2.1) *If $a, b\in K$ then $a<_p b$ in case (p, a, b) R and $a\neq b$.*

Using this notation, our solution of Problem II takes the following form.

THEOREM 2.1. *If a triadic relation R defined on a class K satisfies (α), (β), (t_1), and (t_2), then $(K, \leqq_0)$ is a modular lattice with least element O if and only if (K, R) satisfies (A^O) and (T_5^O) or (A^O) and (T_7^O). When either of these alternatives holds, the relation R is identical with the lattice betweenness of K.*

PROOF. We note first that if (A^0) and (T_8^0) ((A^0) and (T_7^0)) hold for R, the proof of Theorem 1 assures us that $(K, \leqq_0)$ is a lattice. Notice also that $a \leqq_0 b \leqq_0 c$ implies that $(a, b, c)R$ holds by the definition of $\leqq_0$ and t_1. The truth of the second sentence of the theorem is then evident from the following theorem [5, Theorem 10.1].

THEOREM 2.2. *If L is a lattice and R is a triadic relation defined on L, then R is the relation of lattice betweenness of L if and only if the following conditions hold.*

(i) *The relation R satisfies (α), (β), and (t_1).*

(ii) *If $a \leqq b \leqq c$, then we have $(a, b, c)R$.*

(iii) *If the relation abc holds, then in the sublattice generated by a, b, c, the transitivity t_2 holds for R.*

That the modular law holds in $(K, \leqq_0)$ is now implied by t_2 [5, Theorem 9.1]. This completes the proof that (A^0) and (T_8^0) ((A^0) and (T_7^0)) are sufficient. Their necessity follows from Theorem 2.2 and Lemma 1. The proof of Theorem 2.1 is now complete.

3. **A characterization of lattices.** In this final section we shall present our solution of Problem III. The abstract systems which we shall consider consist of a class K of elements $a, b, c, \cdots$ together with a triadic relation R defined on K and satisfying (α), (β), and (t_1). We shall find sufficient conditions that such a system be a lattice under an ordering given by (2.1). In the presence of these conditions, we shall find necessary and sufficient conditions that the postulated relation R coincide with the lattice betweenness of K. The first of these results is given in the following theorem.

THEOREM 3.1. *If a triadic relation R defined on a class K satisfies (α), (β), and (t_1), (A^0) and (T_8^0), then the relation $\leqq_0$ is a partial ordering of K with the lattice property.*

PROOF. The antisymmetry of $\leqq_0$ is a consequence of (β), while the transitive law for $\leqq_0$ may be easily obtained if we notice that identifying d and O in (T_8^0) yields the condition

$$(t_2^0) \qquad (O, b, c)R \cdot (O, x, b)R \rightarrow (O, x, c)R.$$

We may also note that identification of c and d in (T_8^0) yields the condition

$$(3.1) \qquad (O, b, c)R \cdot (b, x, c)R \rightarrow (O, x, c)R.$$

This last condition may be used to replace t_2 in the proof given in Theorem 1 of the sufficiency of the conditions (A^0) and (T_8^0). A careful

examination of our proof will reveal that t_2 was used only to show that the relations $(O,a,a\cup b)R$ and $(O,b,a\cup b)R$ follow from $(O,a\cap b,a\cup b)R$, $(a\cap b,a,a\cup b)R$ and $(a\cap b,b,a\cup b)R$. But it is easy to see that these conclusions follow from the same hypotheses under condition (3.1) as well as under t_2. We can then follow the proof of Theorem 1 from this point until the conclusion that $(K, \leqq_0)$ is a lattice has been reached. The proof of Theorem 3.1 is complete.

REMARK 3. That the sufficient conditions of Theorem 3.1 are not necessary conditions can be seen by the following example. Let the class K consist of the five points O, a, b, c, I, and let the relation R hold for the triples OaI, ObI, OcI, Obc, and bcI, as well as for those then required by (α) and (β). It is easy to verify that t_1 holds in this example while (A^0) fails. But R gives rise through $\leqq_0$ to the simplest non-modular lattice.

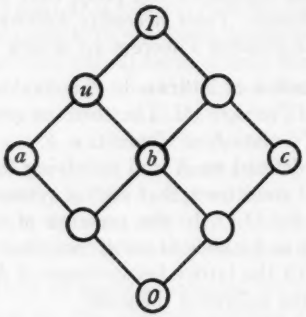

FIG. 1

REMARK 4. Theorem 3.1 leaves us in doubt as to the question of the identity of the triadic relation R and lattice betweenness in K. The following example shows that still further assumptions must be made if we are to obtain this conclusion. Let K consist of the elements of the lattice shown in Figure 1. Define R to be lattice betweenness except that the relation $(a, b, c)R$ is rejected. That R satisfies the conditions (α) and (β) is clear since they hold for lattice betweenness and their conclusions do not involve three distinct elements. To verify that t_1 holds for R we must make sure that we cannot arrange its hypotheses so as to obtain $(a, b, c)R$ in the conclusion. To obtain $(a, b, c)R$ from t_1 would require hypotheses of the form $(d, b, c)R$, $(d, a, b)R$ or of the form $(d, b, a)R$, $(d, c, b)R$. But these sets cannot hold in our example, since if we have $(d, a, b)R$ and $(d, b, c)R$, then $d\neq a$, and $d\cap b\leqq a\leqq d\cup b$, by (1.2). It follows that $d\cup b=u$ or I, con-

trary to $d \cap b \leqq a$. The other set of hypotheses may be treated similarly by interchanging a and c. It is also easy to check that R satisfies (A^0). It remains only to remark that the conclusion of (T_5^0) cannot be $(a, b, c)R$ (since O must appear in the conclusion of (T_5^0)) to assure ourselves that R satisfies all of the conditions of Theorem 4. Thus R generates through $\leqq_0$ the lattice of Figure 1 but R is not the relation of lattice betweenness of this lattice.

To overcome this difficulty we propose the following strengthened condition.

(3.2) *If* $(p \cap x) \cup (x \cap c) = x \cap (p \cup c)$ *and dually, then*

$$(p, b, c)R \cdot (p, d, c)R \cdot (b, x, d)R \rightarrow (p, x, c)R.$$

The proof of Lemma 1 shows that (3.2) holds for lattice betweenness in arbitrary lattices.

We conclude with the following theorem.

THEOREM 3.2. *If a triadic relation R defined on a class K satisfies* (α), (β), (t_1), (T_5^0), *and* (A^0), *then R is the lattice betweenness of the lattice* $(K, \leqq_0)$ *if and only if R satisfies* (3.2).

PROOF. The remark preceding the statement of the theorem disposes of the necessity of (3.2). To establish the sufficiency we rely on the proof of Theorem 10.1 of [5] to assure us that the relation $(a, b, c)R$ implies abc (lattice betweenness). It remains, then, to prove that the relation abc implies the relation $(a, b, c)R$. It was proved in [5] that when abc holds the sublattice generated by a, b, c is distributive. Hence the transitivity (T_5) (see [5]) is available to us *in this sublattice* by virtue of (3.2). Now notice that since $a \cup c \geqq_0 b \geqq_0 a \cap c$ (by (1.2)), we have $a \cup b \cup c = a \cup c$ and dually. The condition (A^0) then gives the relations $(a, a \cup b \cup c, c)R$ and $(a, a \cap b \cap c, c)R$. Clearly $a \cup b \cup c \geqq_0 b \geqq_0 a \cap b \cap c$; from which we obtain $(a \cup b \cup c, b, a \cap b \cap c)R$. The transitivity (T_5^0) then yields $(a, b, c)R$ as desired and the proof is complete.

REFERENCES

1. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications vol 25, 1940.

2. L. M. Blumenthal, *Distance geometries*, University of Missouri Studies vol. 13 (1938) no. 2.

3. V. Glivenko, *Géométrie des systèmes de choses normées*, Amer. J. Math. vol. 58 (1936) pp. 799–828.

4. K. Menger, *Untersuchungen über die allgemeine Metrik*, Math. Ann. vol. 100 (1928) pp. 75–163.

5. Everett Pitcher and M. F. Smiley, *Transitivities of betweenness*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 95–114.

LEHIGH UNIVERSITY

ON A THEOREM OF NEWSOM

H. K. HUGHES

1. **Extension of the theorem.** In 1938, Newsom[1] published a paper containing a theorem regarding the behavior for large values of $|z|$ of the function

$$f(z) = \sum_{n=0}^{\infty} g(n)z^n, \tag{1}$$

radius of convergence equal to ∞. It is assumed that the function $g(w)$, where $w=x+iy$, satisfies the following two conditions:

(a) it is single-valued and analytic in the finite w-plane;

(b) it is such that for all values of x and y, one may write

$$|g(x+iy)| < Ke^{\pi|y|}, \tag{2}$$

where K is a positive constant and k is a positive integer. Under these conditions, according to the theorem, $f(z)$ may be expressed in the form

$$f(z) = \int_{-l-1/2}^{\infty} g(x)[\pm z]^z \frac{\sin k\pi x}{\sin \pi x}\,dx - \sum_{m=0}^{l} \frac{g(-m)}{z^m} + \xi(z, l), \tag{3}$$

where l is any positive integer, where the symbol $[\pm z]^z$ means z^z or $(-z)^z$ according as k is odd or even, respectively, and where if $|\arg [\pm z]| < \pi$, we have $\lim_{|z|\to\infty} z^l \xi(z, l) = 0$ for every value of l.

In the present paper we shall consider the situation when conditions (a) and (b) are made somewhat less restrictive. The theorem which we wish to prove is as follows:

THEOREM. *Let the coefficient $g(n)$ in* (1) *satisfy condition* (a) *except for a singularity at the point $w=w_1$, which is not a negative integer; and let inequality* (2) *be satisfied for all values of $|x|$ and $|y|$ sufficiently large. Then* (3) *continues to hold provided one subtracts from the right member the loop integral*

$$\frac{1}{2i}\int_C \frac{g(w)z^w}{e^{k\pi iw}\sin \pi w}\,dw \tag{4}$$

Received by the editors July 7, 1942.

[1] *On the character of certain entire functions in distant portions of the plane*, Amer. J. Math. vol. 60 (1938) pp. 561–572.

where the loop C surrounds the point w_1, and extends to infinity in any convenient direction lying in either the third or the fourth quadrant.

PROOF. In a concluding remark in his paper, Newsom infers that, under the conditions which we have postulated, the theorem continues to hold provided one subtracts from the right member of (3) a suitable loop integral, which upon examination of the analysis, is seen to be

$$I = \frac{k!\pi^k}{2i}\int_C \frac{P(w, z)}{(\sin \pi w)^{k+1}}\,dw, \tag{5}$$

where[2]

$$P(w, z) = \sum_{j=1}^{k} b_j e^{r_j \pi i w}\int_0^w e^{-r_j \pi i w} g(w)[\pm z]^w dw,$$

$$r = k + 1 - 2j, \qquad 1 \leqq j \leqq k,$$

$$b_j = \frac{(-1)^{j-1}(2\pi i)^{1-k}}{(k-1)!} C_{j-1}^{k-1}, \qquad C_{j-1}^{k-1} = 1 \text{ when } j = 1.$$

We shall show that the integral I in (5) is equivalent to the integral (4). Evidently we may write

$$P(w, z) = \int_0^w \left\{ \sum_{j=1}^{k} b_j e^{r_j \pi i (w-t)} \right\} g(t)[\pm z]^t dt.$$

However, the sum appearing in the above integrand is readily shown to be equal to

$$\frac{1}{\pi^{k-1}(k-1)!}[\sin \pi(w-t)]^{k-1},$$

and this last expression can be written in the form

$$\frac{1}{\pi^{k-1}(k-1)!}\sum_{j=1}^{k}(-1)^{j-1}C_{j-1}^{k-1}(\sin \pi w)^{k-j}(\cos \pi w)^{j-1}(\cos \pi t)^{k-j}(\sin \pi t)^{j-1}.$$

Introducing into (5) the resulting new form for $P(w, z)$, and making simple reductions, we obtain the equation

$$I = \frac{k\pi}{2i}\int_C \left\{ \sum_{j=1}^{k} C_{j-1}^{k-1}(-\cot \pi w)^{j-1}\csc^2 \pi w\, \phi_j(w, z) \right\} dw, \tag{6}$$

where

[2] The form of b_j as given here may be seen to be equivalent to that given by Newsom.

$$\phi_j(w, z) = \int_0^w (\cos \pi t)^{k-j} (\sin \pi t)^{j-1} g(t) [\pm z]^t dt.$$

We may simplify (6) by an integration by parts. Consider first the general term of the sum appearing in the integrand. If we let $u = \phi_j(w, z)$, $dv = \pi(-\cot \pi w)^{j-1} \csc^2 \pi w \, dw$, then, taking account of the constant factors, we obtain as an integral of the general term the expression

$$\frac{k}{2ij} C_{j-1}^{k-1} \left\{ (-\cot \pi w)^j \phi_j(w, z) - \int (-\cot \pi w)^j (\cos \pi w)^{k-j} (\sin \pi w)^{j-1} g(w) [\pm z]^w dw \right\}.$$

Moreover the expression $\phi_j(w, z)$ and the indefinite integral here appearing are both such that they can be evaluated along the loop C. Simplifying the above form and summing with respect to j, we have, as an integral of the entire integrand in (6), the sum

$$\frac{k}{2i} \sum_{j=1}^{k} \frac{(-1)^j}{j} C_{j-1}^{k-1} \left\{ (\cot \pi w)^j \phi_j(w, z) - \int (\cos \pi w)^k \frac{g(w) [\pm z]^w}{\sin \pi w} dw \right\}. \tag{7}$$

If we now agree to let C extend to infinity in a direction of either of the third or fourth quadrants, then at the infinite extremities, the function $\cot \pi w$ has the value i. Consequently, upon evaluating (7) along C and combining properly the definite integrals which arise, we arrive at the result

$$I = \frac{k}{2i} \int_C \left\{ \sum_{j=1}^{k} \frac{(-1)^j}{j} C_{j-1}^{k-1} (a^j b^{k-j} - b^k) \right\} \frac{g(w) [\pm z]^w}{\sin \pi w} dw, \tag{8}$$

where a and b denote $i \sin \pi w$ and $\cos \pi w$, respectively. The sum appearing in the integrand of (8) is easily shown to be equal to $k^{-1}(b-a)^k$, or $k^{-1}e^{-k\pi i w}$. Substituting this quantity for the sum, we have the integral (4).

The above theorem assumes only one singular point of $g(w)$, but the extension to the case when a finite number of singularities occurs is obvious. Moreover, if a singularity is polar in character, then the corresponding loop integral reduces to the residue of the function

$$\frac{1}{2i} \frac{g(w) [\pm z]^w}{e^{k\pi i w} \sin \pi w}$$

at the pole. Furthermore, if a singularity is a negative integer, say $-n$, then the corresponding term in z^{-n} in (3) is to be suppressed and in its place is to be substituted the proper loop integral.

2. **Asymptotic developments.** The theorem of Newsom, together with the extension just established, finds application in the determination of asymptotic developments for large values of $|z|$ of such integral functions as are defined by (1). It is evident that in order to determine completely the asymptotic development of a given function $f(z)$, one must first find such developments for the integral appearing in the right member of (3), and each loop integral (4) that may occur. The first of these questions is a problem about which we shall not concern ourselves here, but we shall discuss briefly a method by which the second question can often be answered.

If the singularity w_1, of $g(w)$, is algebraic in character, then $g(w)$ can be written in the form $g(w)=(w-w_1)^{\theta}\Phi(w)$, where $\Phi(w)$ is analytic at $w=w_1$, $\phi(w_1)\neq 0$, and where θ is a real constant not equal to zero or a positive integer. One of the transformations $w'=\pm(w-w_1)$ will then transform (4) into an integral of the form

$$I(z,\beta)=\frac{i}{2\pi}\int_{C'}F(w)(-w)^{\beta-1}(-[\pm z])^{-w}dw \tag{9}$$

where β is a constant, and where C' is a loop about the origin $w=0$ (the primes having been dropped), and extending to infinity in a direction lying in either the first or the fourth quadrant. Moreover, $F(w)$ is analytic at the origin, and possesses a convergent series development of the form

$$F(w)=\sum_{n=0}^{\infty}c_n(-w)^n,$$

radius of convergence greater than 0. According to a theorem due to Barnes,[3] if $F(w)$ is bounded in the distant right half w-plane, and if the expressions $(-w)^{\beta-1}$ and $(-[\pm z])^{-w}$ are made precise by suitable

[3] For a full statement and proof of the theorem of Barnes, see Ford, *Asymptotic developments of functions defined by Maclaurin series*, Michigan Science Series vol. 11 (1936) p. 16. This book also contains a proof of the theorem of Newsom for the special case in which $k=1$, together with the extension and numerous applications. For proof of the theorem, see chap. 4; for applications, see chaps. 6 and 7. For further applications of the theorems of Ford and Newsom, and of the theorem in the present paper, see C. G. Fry and H. K. Hughes, *Asymptotic developments of certain integral functions*, Duke Math. J. vol. 9 (1942) pp. 791–802.

definitions, then the loop integral $I(z, \beta)$ in (10) is developable asymptotically in the form

$$I(z, \beta) \sim \sum_{n=0}^{\infty} \frac{c_n}{[\log(-[\pm z])]^{\beta+n}\Gamma(1-\beta-n)}.$$

It thus appears that the presence of an algebraic singularity of $g(w)$ presents no serious difficulty.

PURDUE UNIVERSITY

ON SOME FORMULAS INVOLVING THE DIVISOR FUNCTION

HERBERT S. ZUCKERMAN

Viggo Brun[1] has proved the formulas

$$T_1(n) - T_2(n) + T_3(n) - \cdots = -\mu(n), \qquad n > 1, \tag{1}$$

$$h(n) = T_1(n) - (1/2)T_2(n) + (1/3)T_3(n) - \cdots = \begin{cases} 0 \text{ if } n \text{ is not a prime power,} \\ 1/t \text{ if } n = p^t, \ p \text{ a prime;} \end{cases} \tag{2}$$

where $T_l(n)$ is the number of ways that n can be expressed as a product of l factors, each greater than 1. He obtains them as special cases of combinatorial theorems. Pavel Kuhn[2] has also given proofs but it seems that no one has attempted to give elementary number theory proofs of these formulas. It is the purpose of this note to give such proofs and to point out a few other formulas similar to (1) and (2).

All the formulas which we shall prove can be proved very concisely by using the generating function

$$\sum_{n=1}^{\infty} T_l(n)n^{-s} = \{\zeta(s) - 1\}^l,$$

and some simple properties of the zeta-function.[3] Our number theory

Received by the editors June 19, 1942.

[1] Netto, *Lehrbuch der Combinatorik*, 2d edition, 1927, chap. 14, especially pp. 276–277.

[2] *Det Kongelige Norske Videnskabers Selskab*, Forhandlinger, 1939.

[3] Interchanging the order of summation we have $\sum_{n=1}^{\infty}\sum_{l=1}^{\infty}(-1)^{l-1}T_l(n)n^{-s} = \sum_{l=1}^{\infty}(-1)^{l-1}\{\zeta(s)-1\}^l = -\zeta(s)^{-1} = -\sum_{n=1}^{\infty}\mu(n)n^{-s}$, and (1) is obtained by comparing coefficients of n^{-s} in the two members. Similarly, (2) follows from $\sum_{l=1}^{\infty}\sum_{l=1}^{\infty}(-1)^{l-1}l^{-1}T_l(n)n^{-s} = \sum_{l=1}^{\infty}(-1)^{l-1}l^{-1}\{\zeta(s)-1\}^l = \log\zeta(s) = \sum_p \log(1-p^{-s})^{-1} = \sum_p\sum_{l=1}^{\infty} l^{-1}p^{-ls}$.

proofs are the algebraic translations of these proofs.

In order to avoid certain special cases it will be convenient to define

$$T_0(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases} \tag{3}$$

The function $T_1(n)$ has already been defined as the number of ways n can be expressed as a single factor greater than 1, so we have

$$T_1(n) = \begin{cases} 0 & \text{if } n = 1, \\ 1 & \text{if } n > 1. \end{cases} \tag{4}$$

In writing an expression such as the left member of (1) it will be convenient to write it as $\sum_{l=1}^{\infty}(-1)^{l-1}T_l(n)$ but it should be noted that all the terms in the series are zero after a certain point. This will be true of all the series which we write in this way, so we are really dealing only with finite sums.

To prove (1) we first note that from the definition of $T_l(n)$ we have

$$T_{l+1}(n) = \sum_{d|n,\, d\neq n} T_l(d) = \sum_{d|n} T_l(d) - T_l(n), \tag{5}$$

for $l \geqq 0$. We let $f(n) = \sum_{l=0}^{\infty}(-1)^l T_l(n)$ and have, by (5),

$$\sum_{d|n} f(d) = \sum_{l=0}^{\infty}(-1)^l \sum_{d|n} T_l(d) = \sum_{l=0}^{\infty}(-1)^l \{T_l(n) + T_{l+1}(n)\} = T_0(n).$$

Then by the Möbius inversion formula we have

$$f(n) = \sum_{l=0}^{\infty}(-1)^l T_l(n) = \sum_{d|n} T_0(n/d)\mu(d) = \mu(n), \tag{6}$$

where we have used (3). This is equivalent to (1).

This proof has an obvious extension. Putting $g(n) = \sum_{l=1}^{\infty}(-1)^l \cdot (l+1)T_l(n)$ we have, using (5),

$$\begin{aligned} \sum_{d|n} g(d) &= \sum_{l=0}^{\infty}(-1)^l(l+1)\sum_{d|n}^{\infty} T_l(d) \\ &= \sum_{l=0}^{\infty}(-1)^l(l+1)\{T_l(n) + T_{l+1}(n)\} \\ &= \sum_{l=0}^{\infty}(-1)^l T_l(n) = f(n) = \mu(n). \end{aligned}$$

Inverting this we obtain

$$g(n) = \sum_{l=0}^{\infty} (-1)^l (l+1) T_l(n) = \sum_{d|n} \mu(d)\mu(n/d)$$

(7)
$$= \begin{cases} 1 \text{ if } n = 1, \\ (-1)^s 2^s \text{ if } n = p_1 \cdots p_s q_1^2 \cdots q_t^2, \\ 0 \text{ if } n \text{ is divisible by a cube.} \end{cases}$$

In the same way we can prove

$$\sum_{l=0}^{\infty} (-1)^l \frac{(l+1)(l+2)}{2} T_l(n) = \sum_{d\delta|n} \mu(d)\mu(\delta)\mu(n/d\delta)$$

(8)
$$= \begin{cases} 1 \text{ if } n = 1, \\ (-1)^{s+u} 3^{s+t} \text{ if } n = p_1 \cdots p_s q_1^2 \cdots q_t^2 r_1^3 \cdots r_u^3, \\ 0 \text{ if } n \text{ is divisible by a fourth power.} \end{cases}$$

To prove (2) we need another expression for $T_l(n)$. From the definitions we have

(9)
$$T_l(n) = \sum_{r_1 \cdots r_l = n; r_i > 1} 1 = \sum_{r_1 \cdots r_l = n} T_1(r_1) \cdots T_1(r_l),$$

for $l \geqq 1$. Multiplying this by $\log n$, rearranging, and using (9), we have

$$\begin{aligned} T_l(n) \log n &= \sum_{r_1 \cdots r_l = n} T_1(r_1) \cdots T_1(r_l)(\log r_1 + \cdots + \log r_l) \\ &= l \sum_{r_1 \cdots r_l = n} T_1(r_1) \cdots T_1(r_l) \log r_l \\ &= l \sum_{r_l|n} T_1(r_l) \log r_l \sum_{r_1 \cdots r_{l-1} = n/r_l} T_1(r_1) \cdots T_1(r_{l-1}) \\ &= l \sum_{r_l|n} T_1(r_l) \log r_l \cdot T_{l-1}(n/r_l) \\ &= l \sum_{d|n} T_{l-1}(d) T_1(n/d) \log (n/d) \\ &= l \log n \sum_{d|n} T_{l-1}(d) T_1(n/d) - l \sum_{d|n} T_{l-1}(d) T_1(n/d) \log d. \end{aligned}$$

Reducing this by (5) and (4) we have

$$T_l(n) \log n = l T_l(n) \log n - l \sum_{d|n} T_{l-1}(d) T_1(n/d) \log d,$$

which we can write in the form

(10)
$$(1/(l-1)) \sum_{d|n} T_{l-1}(d) T_1(n/d) \log d = (1/l) T_l(n) \log n.$$

If we now let

$$h(n) = \sum_{l=1}^{\infty} \frac{(-1)^{l-1}}{l} T_l(n), \tag{11}$$

we have, by (4) and (10),

$$\begin{aligned}\sum_{d|n} h(d) \log d &= \sum_{l=1}^{\infty} ((-1)^{l-1}/l)\left\{ \sum_{d|n} T_l(d) T_1(n/d) \log d + T_l(n) \log n \right\} \\ &= \sum_{l=1}^{\infty} (-1)^{l-1}\{(1/l) T_l(n) \log n + (1/(l+1)) T_{l+1}(n) \log n\} \\ &= T_1(n) \log n.\end{aligned}$$

Inverting this we obtain

$$\begin{aligned}h(n) \log n &= \sum_{d|n} T_1(n/d)\mu(d) \log n/d \\ &= \log n \sum_{d|n} T_1(n/d)\mu(d) - \sum_{d|n} T_1(n/d)\mu(d) \log d \\ &= \log n\left\{ \sum_{d|n} \mu(d) - \mu(n) \right\} \\ &\quad - \left\{ \sum_{d|n} \mu(d) \log d - \mu(n) \log n \right\} \\ &= - \sum_{d|n} \mu(d) \log d,\end{aligned}$$

where we have used (4) and the fact that $\sum_{d|n}\mu(d) = 0$ for $n > 1$. Now it is easily verified that $-\sum_{d|n}\mu(d) \log d$ takes the value $\log p$ if $n = p^t$ and the value 0 if n is not a prime power.[4] Hence we have

$$h(n) = \begin{cases} 1/t \text{ if } n = p^t, \\ 0 \text{ if } n \text{ is not a prime power,} \end{cases} \tag{12}$$

which is formula (2).

We can combine (6) and (12) to obtain another formula. We first note that, from (9), we have

$$\begin{aligned}T_l(n) &= \sum_{d|n}\left\{ \sum_{r_1 \cdots r_\mu = d} T_1(r_1) \cdots T_1(r_\mu) \sum_{r_{\mu+1} \cdots r_l = n/d} T_1(r_{\mu+1}) \cdots T_1(r_l) \right\} \\ &= \sum_{d|n} T_\mu(d) T_{l-\mu}(n/d), \qquad 0 \leqq \mu \leqq l.\end{aligned} \tag{13}$$

[4] If $n = p^r m$ and p does not divide m we have $-\sum_{d|n}\mu(d) \log d = -\sum_{\delta|m}\sum_{\gamma|p^r} \mu(\delta\gamma) \log (\delta\gamma) = -\sum_{\delta|m}\{\mu(\delta) \log \delta + \mu(\delta p) \log (\delta p)\} = -\sum_{\delta|m}\{\mu(\delta) \log \delta - \mu(\delta) \log (\delta p)\} = \log p \sum_{\delta|m}\mu(\delta)$ and this has the value 0 if $m > 1$, $\log p$ if $m = 1$.

We let

$$
\begin{aligned}
j(n) &= \sum_{l=1}^{\infty}\sum_{\mu=1}^{l}(-1)^{l-1}(1/\mu)T_l(n) \\
&= T_1(n) - (3/2)T_2(n) + (11/6)T_3(n) - (25/12)T_4(n) \\
&\quad + (137/60)T_5(n) - \cdots,
\end{aligned} \tag{14}
$$

and have, by (13),

$$
\begin{aligned}
j(n) &= \sum_{l=1}^{\infty}\sum_{\mu=1}^{l}(-1)^{l-1}(1/\mu)\sum_{d\mid n}T_\mu(d)T_{l-\mu}(n/d) \\
&= \sum_{d\mid n}\sum_{l=1}^{\infty}\sum_{\mu+\nu=l;\mu\geq 1}(-1)^{\mu-1}(1/\mu)T_\mu(d)(-1)^{\nu}T_\nu(n/d) \\
&= \sum_{d\mid n}\left\{\sum_{\mu=1}^{\infty}(-1)^{\mu-1}(1/\mu)T_\mu(d)\sum_{\nu=0}^{\infty}(-1)^{\nu}T_\nu(n/d)\right\} \\
&= \sum_{d\mid n}h(d)f(n/d),
\end{aligned}
$$

where we have used (11) and (6). Then by (6) and (12) we have

$$
\begin{aligned}
j(n) &= \sum_{p^t\mid n}(1/t)\mu(n/p^t) \\
&= \begin{cases} 0 \text{ if } n \text{ is divisible by the squares of two distinct primes,} \\ (-1)^{s+1}/t(t-1) \text{ if } n = p^tq_1\cdots q_s,\ t>1, \\ (-1)^{s-1}s \text{ if } n = q_1\cdots q_s. \end{cases}
\end{aligned} \tag{15}
$$

In quite a similar way we can prove

$$
\begin{aligned}
k(n) &= \sum_{l=2}^{\infty}(-1)^l\sum_{\nu+\mu=l;\nu,\mu\geq 1}(1/\nu\mu)T_l(n) \\
&= T_2(n) - T_3(n) + (11/2)T_4(n) - (5/6)T_5(n) \\
&\quad + (137/180)T_6(n) - \cdots \\
&= \begin{cases} 0 \text{ if } n \text{ has more than two distinct prime factors,} \\ 2/tu \text{ if } n = p^tq^u, \\ \sum_{\nu=1}^{t-1}(1/\nu(t-\nu)) \text{ if } n = p^t. \end{cases}
\end{aligned} \tag{16}
$$

Brun notes that $\sum_{n=1}^{m}T_1(n) = \{m\}$, $\sum_{n=1}^{m}T_2(n) = \sum_{\mu=2}^{\infty}\{m/\mu\}$, $\sum_{n=1}^{m}T_3(n) = \sum_{\mu,\nu=2}^{\infty}\{m/\mu\nu\}$, and so on, where $\{x\}$ denotes the number of integers not less than 2 and not greater than x. Summing (2) he then has

$$H(m) = \sum_{n=1}^{m} h(n) = \{m\} - (1/2)\sum_{\mu=2}^{\infty}\{m/\mu\} + (1/3)\sum_{\mu,\nu=2}^{\infty}\{m/\mu\nu\} - \cdots \tag{17}$$

This formula can be used to determine $h(n)=H(n)-H(n-1)$ and hence, by (2), it gives a method (although impractical) of determining whether n is a prime power. A similar thing can be done with each of our formulas. For example from (14) we have

$$J(m) = \sum_{n=1}^{m} j(n) = \{m\} - (3/2)\sum_{\mu=2}^{\infty}\{m/\mu\} + (11/6)\sum_{\mu,\nu=2}^{\infty}\{m/\mu\nu\} - \cdots.$$

Then $j(n)=J(n)-J(n-1)$, with the aid of (15), determines the number of prime factors of n if n is not divisible by the squares of more than one prime. As a numerical example consider $n=6$. We have

$$J(6) = \{6\} - (3/2)[\{3\} + \{2\}] = 1/2,$$
$$J(5) = \{5\} - (3/2)\{5/2\} = 5/2,$$
$$j(6) = (1/2) - (5/2) = -2,$$

and hence 6 has two distinct prime factors, each entering to the first power.

Brun also points out that from (2) we have

$$H(m) = \sum_{n=1}^{m} h(n) = \sum_{t=1}^{\infty}(1/t)\pi(m^{1/t}).$$

The other formulas do not lead to such simple results but (16) gives

$$K(m) = \sum_{n=1}^{m} k(n) = \sum_{\mu=1}^{\infty}\sum_{p^{\nu}}(1/\mu\nu)\pi((m/p^{\nu})^{1/\mu}).$$

If (10) is put into exponential form it can be stated in a more interesting way. From (10) we have

$$[n^{T_l(n)}]^{1/l} = \left[\prod_{d\mid n;\, d\neq n} d^{T_{l-1}(d)}\right]^{1/(l-1)}$$

The left member can be obtained by writing n in all possible ways as a product of l factors greater than 1, multiplying them together,

and taking the lth root. The right member can be obtained by writing all proper divisors of n in all possible ways as products of $l-1$ factors greater than 1, multiplying them all together, and taking the $(l-1)$th root. Our result is that these two numbers are equal. Thus for $n=24$, $l=3$ we have

$$[(2\cdot 2\cdot 6)(2\cdot 6\cdot 2)(6\cdot 2\cdot 2)(2\cdot 3\cdot 4)(2\cdot 4\cdot 3)(3\cdot 2\cdot 4)(3\cdot 4\cdot 2)(4\cdot 2\cdot 3)(4\cdot 3\cdot 2)]^{1/3} = 2^9\cdot 3^3,$$

$$[(2\cdot 2)\times(2\cdot 3)(3\cdot 2)\times(2\cdot 4)(4\cdot 2)\times(4\cdot 3)(3\cdot 4)\times(6\cdot 2)(2\cdot 6)]^{1/2}=2^9\cdot 3^3.$$

UNIVERSITY OF WASHINGTON

ON THE FOURIER DEVELOPMENTS OF A CERTAIN CLASS OF THETA QUOTIENTS

M. A. BASOCO

1. **Introduction.** In this paper we shall be concerned with the functions $\phi_\alpha^k(z)$ defined by the relation

$$(1) \qquad \phi_\alpha^k(z) \equiv \left\{\frac{d}{dz}\log\vartheta_\alpha(z, q)\right\}^k = \left\{\frac{\vartheta_\alpha'(z, q)}{\vartheta_\alpha(z, q)}\right\}^k, \qquad \alpha = 0, 1, 2, 3,$$

where $\vartheta_\alpha(z, q)$ is a Jacobi theta function and k is a positive integer. In the first place, we shall derive the Fourier developments which represent these functions in a certain strip of the complex plane; it will be seen that the Fourier coefficients of $\phi_\alpha^k(z)$ depend on those of $\phi_\alpha^s(z)$, $s=1, 2, 3, \cdots, k-1$, through a recurrence relation of order k. Secondly, these developments, in conjunction with certain obvious identities, yield, through the method of paraphrase, some general arithmetical formulae of a type first given by Liouville.[1] Indeed, we recover, in a simple manner, some results given without proof by Liouville, which were later proved by Bell[2] through the use of somewhat complex identities involving a certain set of doubly periodic functions of the second kind. One of these results has recently been proved in a strictly elementary, but very ingenious way, by Uspensky.[3] Finally, we indicate some applications of these formulae to the derivation of a certain type of arithmetic and algebraic identities.

2. **The functions $\phi_\alpha^k(z)$.** It should be pointed out that the case $k=1$, is implicit in §§47 and 48 of Jacobi's *Fundamenta nova*.[4] Likewise, the case $k=2$, has been obtained by G. D. Nichols[5] through the use of certain results due to the present writer.[6] The following is a direct derivation of the necessary procedure for the general case; it depends on a straightforward application of contour integration and the theory of residues. It is convenient to treat the two functions $\phi_0^k(z)$ and

Presented to the Society, November 28, 1942; received by the editors June 29, 1942.

[1] J. Math. Pures Appl. (2) vol. 3 (1858) et seq. See, for example, vol. 3 p. 247 (H).

[2] Trans. Amer. Math. Soc. vol. 22 (1921) p. 215 formula (xiv′).

[3] J. V. Uspensky and M. A. Heaslet, *Elementary number theory*, New York, 1939. See chap. 13 p. 462 formula (R).

[4] Jacobi, *Gesammelte Werke*, vol. 1 p. 187.

[5] G. D. Nichols, Tôhoku Math. J. vol. 40 (1935) pp. 252–258.

[6] M. A. Basoco, Bull. Amer. Math. Soc. vol. 38 (1932) pp. 560–568.

$\phi_1^k(z)$ separately; from the results for these functions we may obtain the corresponding ones for $\alpha=2, 3$ by merely replacing z by $z+\pi/2$.

It follows from the properties of the theta functions[7] and Fourier's theorem for analytic functions that

$$(2) \qquad \phi_0^k(z) = \sum_{n=-\infty}^{\infty} A_n^{(k)} e^{-2nis}, \qquad -\Im(\pi\tau/2) < \Im(z) < \Im(\pi\tau/2),$$

where,

$$(3) \qquad A_n^{(k)} = (1/\pi)\int_{-\pi/2}^{\pi/2} \phi_0^k(z) e^{2nis} dz.$$

To evaluate this integral, consider the contour integral $\int_C \phi_0^k(z)e^{2nis}dz$, the contour C being the boundary of the parallelogram with vertices at $z=\pm\pi/2$ and $z=\pm\pi/2+\pi\tau/2$. Cauchy's theorem applied to this integral leads without difficulty to the recurrence relation

$$(4) \qquad A_n^{(k)} - q^{2n}\sum_{j=0}^{k-1} C_{k,j}(-2i)^j A_n^{(k-j)} = 2iR, \qquad q = \exp \pi i\tau,$$

where $C_{k,j}$ is the binomial coefficient $k!/j!(k-j)!$ and R is the residue of the integrand at the pole of order k with affix $z=\pi\tau/2$.

We thus obtain, on placing $k=1, 2, 3$ and computing the necessary residues, the following expressions for the coefficients in the expansions of $\phi_0^k(z)$, for the values of k indicated:

(a) $A_0^{(1)} = 0;\ A_n^{(1)} = -A_{-n}^{(1)} = \dfrac{2iq^n}{1-q^{2n}}, \qquad n \neq 0,$

(b) $A_n^{(2)} = A_{-n}^{(2)} = \dfrac{8q^2}{(1-q^{2n})^2} - \dfrac{4(n-1)}{1-q^{2n}},\ n \neq 0;\ A_0 = -2\sum_{n=1}^{\infty} A_n^{(2)}$

(c) $A_0^{(3)} = 0;\ A_n^{(3)} = -A_{-n}^{(3)} = \dfrac{\vartheta_1'''}{\vartheta_1'}\cdot\dfrac{2iq^n}{1-q^{2n}}$

$$- \frac{2iq^n}{(1-q^{2n})^3}\{(2n^2+6n+3)q^{4n} + (18-4n^2)q^{2n} + (2n^2-6n+3)\}.$$

With respect to the function $\phi_1^k(z)$ the process used in the preceding must be modified slightly on account of the presence of a singularity (pole of order k) at the origin. Define the function $\psi_1^{(k)}(z)$ to be of the form

[7] See Whittaker and Watson, *Modern analysis*, chap. 21.

$$\psi_1^{(k)}(z) \equiv \phi_1^k(z) - T^{(k)}(z), \tag{5}$$

where $T^{(k)}(z)$ is a suitable function of $\sin z$ and $\cos z$, having the period π and such that the principal part of its Laurent expansion about the origin coincides with that of $\phi_1^k(z)$. The function $\psi_1^{(k)}(z)$ is, therefore, analytic along the real axis. It is found that, for the cases considered (that is, for $k=1, 2, 3$), the functions $T^{(k)}(z)$ are of the form:

$$T^{(1)}(z) = \cot z; T^{(2)}(z) = \cot^2 z; T^{(3)}(z) = (1+\vartheta_1'''/\vartheta_1')\cot z + \cot^3 z.$$

We may now write

$$\psi_1^{(k)}(z) = \sum_{n=1}^{\infty} B_n^{(k)}(e^{2niz} - e^{-2niz}), \text{ if } k \text{ is odd},$$

$$\psi_1^{(k)}(z) = B_0^{(k)} + \sum_{n=1}^{\infty} B_n^{(k)}(e^{2niz} + e^{-2niz}), \text{ if } k \text{ is even}.$$

These expansions are valid in the strip defined by the inequalities $-\Im(\pi\tau) < \Im(z) < \Im(\pi\tau)$. In either case, we have,

$$B_n^{(k)} = (1/\pi)\int_{-\pi/2}^{\pi/2} \psi_1^{(k)}(z)e^{-2niz}\,dz, \qquad n \neq 0. \tag{6}$$

The value of $B_0^{(k)}$ can be obtained from the relation

$$B_0^{(k)} = -2\sum_{n=1}^{\infty} B_n^{(k)} + \lim_{z\to 0} \psi_1^{(k)}(z). \tag{7}$$

To calculate the integral in (6), consider the contour integral $\int_C \psi_1^{(k)}(z)e^{-2niz}dz$ where C is the boundary of the parallelogram with vertices at $z=\pm\pi/2$ and $z=\pm\pi/2-(2s+1)\pi\tau/2$, where s is an arbitrary positive integer. This contour contains precisely s poles of the function $\psi_1^{(k)}(z)$, these being located at the points with affixes $z=-r\pi\tau$, $r=1, 2, 3, \cdots, s$. On account of the properties of the functions involved we may pass to the limit as s tends to infinity, thus obtaining

$$B_n^{(k)} = -2i\sum_{s=1}^{\infty} R_s$$

where R_s is the residue of $\psi^{(k)}(z)e^{-2niz}$ at the poles of order k, $z=-s\pi\tau$.

It is found in this manner that

$$B_0^{(1)} = 0;\ B_n^{(1)} = \frac{-2iq^{2n}}{1-q^{2n}}, \qquad n \neq 0. \tag{d}$$

(e) $B_0^{(2)} = 8\sum_{n=1}^{\infty} \frac{nq^{2n}}{1-q^{2n}}$; $B_n^{(2)} = \frac{8q^{2n}}{(1-q^{2n})^2} - \frac{4nq^{2n}}{1-q^{2n}}$, $n \neq 0$.

(f) $B_0^{(3)} = 0$; $B_n^{(3)} = \frac{\vartheta_1'''}{\vartheta_1'} \cdot \frac{4q^{2n}}{1-q^{2n}}$

$$- \frac{8q^{2n}}{(1-q^{2n})^3}\{n^2q^{4n} - (2n^2 - 6n - 6)q^{2n} + (n^2 - 6n + 6)\}, \quad n \neq 0.$$

On substituting the values of the coefficients given in (a), $\cdots$, (f), into (2) and (5) and transforming the resulting series into an arithmetical form, we find the following:

(I) $\phi_0(z) = 4\sum_{(n)} q^n\{\sum \sin 2tz\}$, $n = 1, 2, 3, 4, \cdots$,

(II) $\phi_1(z) = \cot z + 4\sum_{(n)} q^{2n}\{\sum \sin 2dz\}$,

(III) $\phi_2(z) = -\tan z + 4\sum_{(n)} q^{2n}\{\sum (-1)^d \sin 2dz\}$,

(IV) $\phi_3(z) = 4\sum_{(n)} (-1)^n q^n\{\sum \sin 2tz\}$,

(V) $\phi_0^2(z) = 8\sum_{(n)} q^n \sigma(n) + 8\sum_{(n)} q^n\{\sum (\tau - t) \cos 2tz\}$,

(VI) $\phi_1^2(z) = \cot^2 z + 8\sum_{(n)} q^{2n} \zeta_1(n) + 8\sum_{(n)} q^{2n}\{\sum (2\delta - d) \cos 2dz\}$,

(VII) $\phi_2^2(z) = \tan^2 z + 8\sum_{(n)} q^{2n} \zeta_1(n)$

$$- 8\sum_{(n)} q^{2n}\{\sum (-1)^d (2\delta - d) \cos 2dz\},$$

(VIII) $\phi_3^2(z) = 8\sum_{(n)} (-1)^n q^n \sigma(n) + 8\sum_{(n)} (-1)^n q^n\{\sum (\tau - t) \cos 2tz\}$,

(IX) $\phi_0^3(z) = (\vartheta_1'''/\vartheta_1')\phi_0(z) - 4\sum_{(n)} q^n\{\sum (3\tau^2 - 6t\tau + 2t^2) \sin 2tz\}$,

(X) $\phi_1^3(z) = \cot z + \cot^3 z + (\vartheta_1'''/\vartheta_1')\phi_1(z)$

$$- 8\sum_{(n)} q^{2n}\{\sum (d^2 - 6d\delta + 6\delta^2) \sin 2dz\},$$

(XI) $\phi_2^3(z) = -\tan z - \tan^3 z + (\vartheta_1'''/\vartheta_1')\phi_2(z)$

$$- 8\sum_{(n)} q^{2n}\{\sum (-1)^d (d^2 - 6d\delta + 6\delta^2) \sin 2dz\},$$

$$\phi_3^3(z) = (\vartheta_1'''/\vartheta_1')\phi_3(z) - 4\sum_{(n)}(-1)^n q^n\{\sum(3\tau^2 - 6t\tau + 2t^2)\sin 2tz\},$$

(XII)

$$(\vartheta_1'''/\vartheta_1') = -1 + 24\sum_{(n)} q^{2n}\zeta_1(n).$$

In these expansions, the notation is as usual: The inner sigma refers to the divisors d, δ, t, τ of the positive integer n, τ being odd, and

$$\sigma(n) = \psi_1(n) - \zeta_1'(n)$$

where,

$\psi_1(n)$ = sum of the divisors of n whose conjugates are odd,
$\zeta_1'(n)$ = sum of the odd divisors of n, and
$\zeta_1(n)$ = sum of all the divisors of n.

3. **Paraphrases.** The obvious identities

$$\phi_a(z)\cdot\phi_a(z) = \phi_a^2(z), \qquad a = 0, 1, 2, 3, \tag{8}$$

in conjunction with the expansions (I) to (VIII) of the preceding section yield, through the method of paraphrase,[8] the following arithmetical theorems:

THEOREM (α). *Let $F(x)$ be an even, single-valued function which is well defined for all integral values of the argument including zero, but is otherwise arbitrary; let n be any positive integer. Then,*

$$\sum_{(k)}\{F(t' - t'') - F(t' + t'')\} = \sum_{(l)}(\tau - t)\{F(t) - F(0)\},$$

where (k) and (l) refer to the following partitions of n in positive integers:

$$(k)\quad n = t'\tau' + t''\tau'', \qquad (l)\quad n = t\tau, \qquad \tau, \tau', \tau'' \text{ odd.}$$

This is the arithmetical equivalent of (8) with $a=0$. It is of interest to note that a strictly elementary proof of this theorem has recently been given by Uspensky.[3]

THEOREM (β). *Let $F(x)$ and the integer n be as in Theorem (α). Then,*

$$\sum_{(i)}\{F(d' - d'') - F(d' + d'')\} = \{\zeta_1(n) - \zeta_0(n)\}F(0) + \sum_{(j)}(2\delta - d - 1)F(d) - 2\sum_{(f)}\{F(1) + F(2) + F(3) + \cdots + F(d-1)\}$$

[8] E. T. Bell, Trans. Amer. Math. Soc. vol. 22 (1921) pp. 1–30; 198–219.

where (i), (j) *refer to the following partitions of* n *in positive integers*:

$$(i)\quad n = d'\delta' + d''\delta'', \qquad (j)\quad n = d\delta.$$

$\zeta_0(n)$ *represents the number of positive integral divisors of* n, *while* $\zeta_1(n)$ *represents their sum.*

This theorem was first given by Liouville,[9] who stated it without giving a proof; it is conjectured that he obtained this as well as other similar results through the aid of the theory of elliptic functions. At any rate, this theorem was later rediscovered by Bell (loc. cit.) as a special case of a more general result derived from one of the addition theorems of the theta functions and involving a certain set of doubly periodic functions of the second kind. We have obtained this result as the equivalent of the identity (8) with $a=1$.

THEOREM (γ). *Let* $F(x)$, n, *and the partitions* (i) *and* (j) *be as in the preceding theorem. Then,*

$$\sum_{(i)} (-1)^{d'+d''}\{F(d'-d'') - F(d'+d'')\}$$
$$= \{\zeta_1(n) - \zeta_0(n)\}F(0) - \sum_{(j)} (-1)^d(2\delta - d + 1)F(d)$$
$$+ 2\sum_{(j)} \{F(1) - F(2) + F(3) - \cdots + (-1)^d F(d-1)\}.$$

This theorem is the arithmetical equivalent of identity (8) with $a=2$. It is, however, not essentially distinct from Theorem (β); for either theorem may be obtained from the other by replacing $F(x)$ by $(-1)^x F(x)$. Finally it should be noted that the case $a=3$ merely yields the result stated in Theorem (α).

As might be expected, those theorems analogous to the preceding which may be deduced from identities involving the functions $\phi_a^3(z)$ are considerably more complicated; it will perhaps suffice, in the present instance, to state merely the following two results by way of samples.

THEOREM (δ). *Let* $G(x)$ *be an odd, single-valued function, which is well defined for all integral values of the argument, but is otherwise arbitrary; let* n *be any positive integer. Then,*

$$4\sum_{(i)} \{G(t_1 - t_2 + t_3) + G(t_1 + t_2 - t_3) + G(t_2 + t_3 - t_1) - G(t_1 + t_2 + t_3)\}$$
$$= 24\sum_{(j)} \{\zeta_1(n_1) \sum G(t_2)\} - \sum_{(k)} (3\tau^2 - 6t\tau + 2t^2 + 1)G(t),$$

[9] Liouville, loc. cit.

where (i), (j) and (k) refer to the following partitions of n in positive integers:

$$(i)\quad n = t_1\tau_1 + t_2\tau_2 + t_3\tau_3 \qquad \tau_1, \tau_2, \tau_3 \text{ all odd},$$

$$(j)\quad n = 2n_1 + n_2, \qquad n_2 = t_2\tau_2, \qquad \tau_2 \text{ odd},$$

$$(k)\quad n = t\tau, \qquad \tau \text{ odd}.$$

This theorem is the arithmetical equivalent of the otherwise apparently trivial identity:

$$\phi_0(z)\cdot\phi_0(z)\cdot\phi_0(z) \equiv \phi_0^3(z).$$

THEOREM (ϵ). *Let $F(x)$, n, and the partitions (i) and (j) be as in Theorem (β). Then,*

$$6\sum_{(i)} d'(d'' - 2\delta'')\{F(d' + d'') + F(d' - d'')\}$$
$$= 6(n\zeta_0(n) - \zeta_2(n))F(0) + 12\sum_j (\delta - d)\left\{\sum_{r=1}^{d}\sum_{s=1}^{r-1} F(s)\right\}$$
$$+ \sum_{(j)} (d^3 - 6d^2\delta + 6d\delta^2 - d)F(d).$$

This theorem results on paraphrasing the identity:

$$\frac{d}{dz}\phi_1^3(z) \equiv 3\phi_1^2(z)\frac{d}{dz}\phi_1(z).$$

4. **An application.** If in Theorem (α) we place $F(x) = x^2$, the following may be deduced

$$4\sum_{k=1}^{n-1}\psi_1(k)\psi_1(n-k) = \psi_3(n) - n\psi_1(n), \tag{9}$$

where $\psi_r(n)$ = sum of the rth powers of the divisors of n whose conjugates are odd. In particular, if n is a prime p, we obtain

$$4\sum_{k=1}^{p-1}\psi_1(k)\psi_1(p-k) = (p-1)(p^2-1).$$

It also follows from (9) and the easily verified relationships:

$$\sum_{n=1}^{\infty}\frac{nx^n}{1-x^{2n}} = \sum_{n=1}^{\infty}\psi_1(n)x^n, \quad \sum_{n=1}^{\infty}\frac{n^3x^n}{1-x^{2n}} = \sum_{n=1}^{\infty}\psi_3(n)x^n,$$
$$\sum_{n=1}^{\infty}\frac{n^2(1+x^{2n})x^n}{(1-x^{2n})^2} = \sum_{n=1}^{\infty}n\psi_1(n)x^n,$$

that,

$$4\left\{\sum_{n=1}^{\infty}\frac{nx^n}{1-x^{2n}}\right\}^2=\sum_{n=1}^{\infty}\frac{n^3x^n}{1-x^{2n}}-\sum_{n=1}^{\infty}\frac{n^2x^n}{(1-x^{2n})^2}-\sum_{n=1}^{\infty}\frac{n^2x^{3n}}{(1-x^{2n})^2}. \tag{10}$$

Again, if in Theorem (β) we place $F(x)=x^2$, we deduce

$$12\sum_{k=1}^{n-1}\zeta_1(k)\zeta_1(n-k)=5\zeta_3(n)-(6n-1)\zeta_1(n), \tag{11}$$

which in turn implies the following development,

$$12\left\{\sum_{n=1}^{\infty}\frac{nx^n}{1-x^n}\right\}^2=\sum_{n=1}^{\infty}(5\zeta_3(n)-(6n-1)\zeta_1(n))x^n. \tag{12}$$

Finally, in Theorem (δ) place $G(x)=x$ to deduce

$$24\sum\zeta_1(n_1)\psi_1(n_2)=2\psi_3(n)-(6n-1)\psi_1(n)+3n\zeta_1'(n), \quad n=2n_1+n_2,\ n_1,\ n_2>0. \tag{13}$$

This result may be interpreted so as to yield the following relation:

$$24\left\{\sum_{n=1}^{\infty}\frac{nx^{2n}}{1-x^{2n}}\right\}\cdot\left\{\sum_{n=1}^{\infty}\frac{nx^n}{1-x^{2n}}\right\}=2\sum_{n=1}^{\infty}\frac{n^3x^n}{1-x^{2n}}+\sum_{n=1}^{\infty}\frac{nx^n}{1-x^{2n}}-6\sum_{n=1}^{\infty}\frac{n^2(1+x^{2n})x^n}{(1-x^{2n})^2}+3\sum_{n=1}^{\infty}\frac{(2n-1)^2x^{2n-1}}{(1-x^{2n-1})^2}. \tag{14}$$

Results analogous to the preceding, but developed from an entirely different point of view, have been obtained by Glaisher.[10] Here, such results appear as very special consequences of the general theorems stated in §3. It is clear, of course, that similar results may be obtained through the full use of the expansions here given.

THE UNIVERSITY OF NEBRASKA

[10] J. W. L. Glaisher, Messenger of Mathematics nos. 166, 169 (1885).

THE BETTI GROUPS OF THE PRODUCT OF TWO NORMAL SPACES

C. E. CLARK

1. **Introduction.** Let R_1 and R_2 be homeomorphic to open sets of normal spaces. Furthermore, let each of R_1 and R_2 contain infinitely many points. For these spaces Alexandroff has defined inner Betti groups.[1] In this paper the inner Betti groups of the topological product $R_1 \otimes R_2$ are studied.

Let $\bar{u}^r \in B^r(R_1)$, where B^r denotes the r-dimensional inner cohomology group with the integers as coefficient domain, and let $\bar{u}^\rho \in B^\rho(R_2)$. To these two elements there corresponds an element $\bar{u}^r \times \bar{u}^\rho \in B^{r+\rho}(R_1 \times R_2)$. All such elements with $r+\rho=n$ generate a subgroup $B_1^n(R_1 \times R_2) \subset B^n(R_1 \times R_2)$. We characterize this subgroup (Theorem 3). In addition, we characterize the factor-group $B_2^n(R_1 \times R_2) = B^n(R_1 \times R_2)/B_1^n(R_1 \times R_2)$ (Theorem 4). In doing so we show that if $\bar{u}^s \in B^s(R_1)$ is of order $\bar{e}^s \neq 0$, and if $\bar{u}^\sigma \in B^\sigma(R_2)$ is of order $\bar{e}^\sigma \neq 0$, then to these two elements there corresponds an element $(\bar{u}^s, \bar{u}^\sigma) \in B_2^{s+\sigma-1}(R_1 \times R_2)$.

To prove these results we employ Alexandroff's second definition of the inner Betti groups which uses barycentric subdivisions of coverings.[2] Furthermore, Freudenthal's simplicial division of the product of two simplexes is used.[3]

2. **The groups $B_i^n(K^a \times K^\alpha \bmod C^a C^\alpha)$, $i=1, 2$.** In this section we state without proof some facts about products of complexes which are consequences of [3]. Let K^a and K^α be finite complexes with subcomplexes C^a and C^α, respectively. Let B^n denote the n-dimensional integral cohomology group. To $u^r \in B^r(K^a \bmod C^a)$ and $u^\rho \in B^\rho(K^\alpha \bmod C^\alpha)$ there corresponds the product of these cohomology classes $u^r \times u^\rho \in B^{r+\rho}(K^a \times K^\alpha \bmod C^a C^\alpha)$. We define $B_1^n(K^a \times K^\alpha \bmod C^a C^\alpha)$ to be the subgroup of $B^n(K^a \times K^\alpha \bmod C^a C^\alpha)$ generated by—we could say consisting of—all $u^r \times u^\rho$ with $r+\rho=n$. Let e^r and e^ρ be the orders of u^r and u^ρ, respectively, with the understanding that $e=0$ when u is free. Let (a, b) denote the greatest common divisor of a and b with the understanding that $(a, 0)=a$.

Received by the editors July 2, 1942.

[1] See [1, 7.2] (Numbers in brackets refer to references at end of paper).

[2] See [1, 9.22 and 9.4].

[3] See [2].

THEOREM 1. *A complete set of relations for the generators* $u^r \times u^\rho$ *of* $B_1^n(K^a \times K^\alpha \bmod C^{a\alpha})$ *is given by*

$$(e^r, e^\rho)(u^r \times u^\rho) = 0 \tag{1}$$

and

$$\sum p_i u_i^r + \sum q_j u_j^\rho = \sum p_i q_j (u_i^r \times u_j^\rho) \tag{2}$$

where the p*'s and* q*'s are integers.*

We define $B_2^n(K^a \times K^\alpha \bmod C^{a\alpha}) = B^n(K^a \times K^\alpha \bmod C^{a\alpha})/B_1^n(K^a \times K^\alpha \bmod C^{a\alpha})$. This factor-group will be characterized in a way suitable for our later discussion.

Let $u^s \in B^s(K^a \bmod C^a)$ be of order $e^s \neq 0$ and $u^\sigma \in B^\sigma(K^\alpha \bmod C^\alpha)$ be of order $e^\sigma \neq 0$. Throughout the paper the superscripts s and σ will be used only for elements of order different from 0, while superscripts r and ρ indicate that no restriction is made on the orders of the elements. We shall associate with u^s and u^σ an element $(u^s, u^\sigma) \in B_2^{s+\sigma-1}(K^a \times K^\alpha \bmod C^{a\alpha})$. To do so let $z^s \in u^s$, $\dot{f}^{s-1} = e^s z^s$ where the dot denotes the coboundary operator, $z^\sigma \in u^\sigma$, and $\dot{f}^{\sigma-1} = e^\sigma z^\sigma$. For any cocycle z let $H(z)$ denote the cohomology class with $z \in H(z)$. We define

$$H\left\{\frac{1}{(e^s, e^\sigma)}(f^{s-1} \times f^{\sigma-1})\cdot\right\} \in (u^s, u^\sigma). \tag{3}$$

We can show that (u^s, u^σ) is independent of the choice of z's and f's.

THEOREM 2. *The cosets* (u^s, u^σ), $s+\sigma-1=n$, *generate* $B_2^n(K^a \times K^\alpha \bmod C^{a\alpha})$, *and a complete set of relations for these generators consists of*

$$(e^s, e^\sigma)(u^s, u^\sigma) = 0 \tag{4}$$

and

$$\left(\sum \mathring{p}_i \mathring{u}_i,\ \sum \mathring{p}_j \mathring{u}_j\right) = \frac{\mathring{m}\mathring{m}}{(m^s, m^\sigma)} \sum \frac{\mathring{p}_i \mathring{p}_j (\mathring{e}_i, \mathring{e}_j)}{e_i^s e_j^\sigma} (\mathring{u}_i, \mathring{u}_j) \tag{5}$$

where we consider only those sums $\sum p_i^s u_i^s$ *and* $\sum p_j^\sigma u_j^\sigma$ *with the following properties: the* p*'s are integers, and the orders* m^s *and* m^σ *of* $\sum p_i^s u_i^s$ *and* $\sum p_j^\sigma u_j^\sigma$ *are the least common multiples of the orders of the terms* $p_i^s u_i^s$, $i = 1, 2, \cdots$, *and* $p_j^\sigma u_j^\sigma$, $j = 1, 2, \cdots$, *respectively.*

3. **Coverings and projections.** We first associate with cofinal systems of coverings[4] of R_1 and R_2 a cofinal system for $R_1 \times R_2$ (our cover-

[4] See [1, 3.5].

ings are finite and consist of open sets). With Alexandroff we say that a covering is multiplicative[5] if the intersection of any two sets of the covering is also a set of the covering. Let $\{\Omega^a\}$ and $\{\Omega^\alpha\}$ be cofinal systems of multiplicative coverings of R_1 and R_2, respectively. Let $\Omega^{a\alpha}=\Omega^a\times\Omega^\alpha$. Let $\Omega^{b\beta}$ follow $\Omega^{a\alpha}$ if both Ω^b follows Ω^a and Ω^β follows Ω^α. It is easily seen that $\{\Omega^{a\alpha}\}$ is a cofinal system of multiplicative coverings of $R_1\times R_2$.

If Ω^b follows Ω^a, let S_a^b be the canonical projection[6] of Ω^b into Ω^a; that is, if $M^b\in\Omega^b$, then $S_a^bM^b$ is the smallest open set of Ω^a such that $M^b\supset S_a^bM^b$. Similarly we have the canonical projection S_α^β of Ω^β into Ω^α if Ω^β follows Ω^α. With these two projections we shall associate the canonical projection $S_{a\alpha}^{b\beta}$ of $\Omega^{b\beta}$ into $\Omega^{a\alpha}$. Let $M^b\in\Omega^b$ and $M^\beta\in\Omega^\beta$. It is easily seen that

$$S_{a\alpha}^{b\beta}M^b\times M^\beta=S_a^bM^b\times S_\alpha^\beta M^\beta. \tag{6}$$

By the barycentric subdivision of a covering[7] is meant that subcomplex of the nerve of the covering made up of the simplexes whose vertices are associated with decreasing sequences of open sets of the covering. Let K^a, K^α, and $K^{a\alpha}$ be the barycentric subdivisions of Ω^a, Ω^α, and $\Omega^{a\alpha}$. We shall show that $K^{a\alpha}$ is a simplicial subdivision of $K^a\times K^\alpha$.

First, let the r-simplex t^r of K^a be associated with the following sets of Ω^a:

$$M_0^a\supset\cdots\supset M_r^a. \tag{7}$$

Similarly, let t^ρ of K^α be associated with

$$M_0^\alpha\supset\cdots\supset M_\rho^\alpha. \tag{8}$$

Corresponding to these two simplexes there is a set of $(r+\rho)$-simplexes of $K^{a\alpha}$ in (1-1)-correspondence with the sequences

$$M_0^a\times M_0^\alpha\supset\cdots\supset M_i^a\times M_j^\alpha\supset M_k^a\times M_l^\alpha\supset\cdots\supset M_r^a\times M_\rho^\alpha \tag{9}$$

where $k\geqq i$, $l\geqq j$, and $k+l=i+j+1$. But Freudenthal has shown[8] that this set of simplexes forms a simplicial division of $t^r\times t^\rho$. Furthermore, the cells $t^r\times t^\rho$ determined by (9) form a cell complex that is isomorphic with $K^a\times K^\alpha$.

---

[5] See [1, 9.2].
[6] See [1, 9.3].
[7] See [1, 9.22].
[8] See [2].

Each of the complexes K^a, K^α, and $K^{a\alpha}$ has a special subcomplex[9] C^a, C^α, and $C^{a\alpha}$ made up of the simplexes all of whose vertices are associated with open sets whose closures are not bicompact. We shall show that $C^a \times K^\alpha + K^a \times C^\alpha \subset C^{a\alpha}$. Indeed, suppose the closure of each set of (7) is not bicompact. Then the closure of each set of any sequence of (9) is not bicompact (because the continuous image of a bicompact set is bicompact, and a factor of a product of two sets is a continuous image of the product). The same argument holds for a simplex of C^α, and the above inclusion holds.

Next suppose t^r is not in C^a. This means that $\overline{M}^a_r$, the closure of the final term of (7), is bicompact. Also suppose t^ρ is not in C^α. This means that $\overline{M}^\alpha_\rho$, from (8), is bicompact. Since the product of the closures of two sets is the closure of the product, and since the product of two bicompact sets is bicompact, the last term of (9) has a bicompact closure. This means that $t^r \times t^\rho$ is not in $C^{a\alpha}$. This with the above inclusion means that $C^a \times K^\alpha + K^a \times C^\alpha = C^{a\alpha}$. We have proved this lemma.

LEMMA. *The complex $K^{a\alpha}$ is a simplicial division of $K^a \times K^\alpha$, and $C^{a\alpha}$ is the division of $C^a \times K^\alpha + K^a \times C^\alpha$.*

The canonical projections S^b_a, S^β_α, and $S^{b\beta}_{a\alpha}$ determine the usual homomorphic mappings ρ^b_a, ρ^β_α, and $\rho^{b\beta}_{a\alpha}$ of the chains of K^b mod C^b, and so on, into the chains of K^a mod C^a, and so on.[10] We shall show that

$$\rho^{b\beta}_{a\alpha}(t^r \times t^\rho) = \rho^b_a t^r \times \rho^\beta_\alpha t^\rho. \tag{10}$$

In (7) replace the superscript a by b and let the resulting sequence denote the open sets associated with t^r. In the same way we replace α by β in (8) and (9) to obtain sequences associated with t^ρ and $t^r \times t^\rho$. From (6) and (9) we see that the left side of (10) is associated with the sequences $S^b_a M^b_0 \times S^\beta_\alpha M^\beta_0 \supset \cdots$. But these sequences are associated with $\rho^b_a t^r \times \rho^\beta_\alpha t^\rho$ because $\rho^b_a t^r$ is associated with $S^b_a M^b_0 \supset \cdots$ and $\rho^\beta_\alpha t^\rho$ is associated with $S^\beta_\alpha M^\beta_0 \supset \cdots$.

The canonical projections S^b_a, S^β_α, and $S^{b\beta}_{a\alpha}$ determine the usual homomorphic mappings σ^a_b, σ^α_β, and $\sigma^{a\alpha}_{b\beta}$ of the chains of K^a mod C^a, and so on, into the chains of K^b mod C^b, and so on.[10] From (10) we have

$$\sigma^{a\alpha}_{b\beta}(f^a \times f^\alpha) = \sigma^a_b f^a \times \sigma^\alpha_\beta f^\alpha. \tag{11}$$

Furthermore, these σ's determine homomorphic mappings π^a_b of

[9] See [1, 9.1].

[10] See [1, 2.2].

$B^n(K^a \bmod C^a)$ into $B^n(K^b \bmod C^b)$, π_β^α of $B^n(K^\alpha \bmod C^\alpha)$ into $B^n(K^\beta \bmod C^\beta)$, and in view of the lemma $\pi_{b\beta}^{a\alpha}$ of $B^n(K^a \times K^\alpha \bmod C^{a\alpha})$ into $B^n(K^b \times K^\beta \bmod C^{b\beta})$.[11] From (11) we have

$$\pi_{b\beta}^{a\alpha}(u^a \times u^\alpha) = \pi_b^a u^a \times \pi_\beta^\alpha u^\alpha. \tag{12}$$

4. **The inner Betti groups of $R_1 \times R_2$.** Let $B^n(R_1)$, $B^n(R_2)$, and $B^n(R_1 \times R_2)$ be the inner Betti groups defined in §1. Let $\bar{u}^r \in B^r(R_1)$ be of order $\bar{e}^r$, and $\bar{u}^\rho \in B^\rho(R_2)$ be of order $\bar{e}^\rho$. We shall define $\bar{u}^r \times \bar{u}^\rho \in B^{r+\rho}(R_1 \times R_2)$. Let $u^{r\alpha} \in \bar{u}^r$ and $u^{\rho\alpha} \in \bar{u}^\rho$ where a and α are indices of coverings. From (12) we have $\pi_{b\beta}^{a\alpha}(u^{ra} \times u^{\rho\alpha}) = \pi_b^a u^{ra} \times \pi_\beta^\alpha u^{\rho\alpha}$. Hence we can define

$$u^{ra} \times u^{\rho\alpha} \in \bar{u}^r \times \bar{u}^\rho \tag{13}$$

and determine a bundle independent of the choice of coverings.

A consequence of (2) is

$$\sum p_i \bar{u}_i^r \times \sum q_j \bar{u}_j^\rho = \sum p_i q_j (\bar{u}_i^r \times \bar{u}_j^\rho). \tag{14}$$

It is possible to find coverings with indices b and β so that u^{rb} and $u^{\rho\beta}$ are of order $\bar{e}^r$ and $\bar{e}^\rho$, respectively, (indeed, $\bar{e}^r u^{ra}$ is in the zero bundle, and any cohomology class of the zero bundle can be projected into a zero cohomology class; furthermore, u^{rb} cannot be of order less than $\bar{e}^r$). This with (1) gives

$$(\bar{e}^r, \bar{e}^\rho)(\bar{u}^r \times \bar{u}^\rho) = 0. \tag{15}$$

Let $B_1^n(R_1 \times R_2)$ be the subgroup of $B^n(R_1 \times R_2)$ generated by the bundles (13) with $r+\rho = n$.

THEOREM 3. *Relations* (14) *and* (15) *are a complete set for the generators* (13) *of* $B_1^n(R_1 \times R_2)$.

PROOF. Let the finite sum $\sum p_{ij} \bar{u}_i^r \times \bar{u}_j^\rho = 0$, $r+\rho = n$, but r and ρ can vary from term to term. For some coverings with indices a and α we have $\sum p_{ij}(u^{ra} \times u^{\rho\alpha}) = 0$. Theorem 1 proves that this latter relation is a consequence of (1) and (2). Hence the original relation is a consequence of (14) and (15) q.e.d.

Equation (12) shows that $\pi_{b\beta}^{a\alpha} B_1^n(K^a \times K^\alpha \bmod C^{a\alpha}) \subset B_1^n(K^b \times K^\beta \bmod C^{b\beta})$. Hence there is a natural homomorphism $\Pi_{b\beta}^{a\alpha}$ of $B_2^n(K^a \times K^\alpha \bmod C^{a\alpha})$ into $B_2^n(K^b \times K^\beta \bmod C^{b\beta})$. This natural homomorphism is defined as follows. If $\phi^{a\alpha} u^{a\alpha}$ is the coset of $B_2^n(K^a \times K^\alpha \bmod C^{a\alpha})$ that contains the cohomology class $u^{a\alpha} \in B^n(K^a \times K^\alpha \bmod C^{a\alpha})$, then

[11] See [1, 2.33].

$$\Pi^{\alpha\alpha}_{b\beta}\phi^{a\alpha}u^{a\alpha} = \phi^{b\beta}\pi_{b\beta}u^{a\alpha} \tag{16}$$

We shall show that $\Pi^{b\beta}_{c\gamma}\Pi^{a\alpha}_{b\beta} = \Pi^{a\alpha}_{c\gamma}$. Indeed, using (16) and the fact that $\pi^{b\beta}_{c\gamma}\pi^{a\alpha}_{b\beta} = \pi^{a\alpha}_{c\gamma}$, we have $\Pi^{a\alpha}_{c\gamma}\phi^{a\alpha}u^{a\alpha} = \phi^{c\gamma}\pi^{a\alpha}_{c\gamma}u^{a\alpha} = \phi^{c\gamma}\pi^{b\beta}_{c\gamma}\pi^{a\alpha}_{b\beta}u^{a\alpha} = \Pi^{b\beta}_{c\gamma}\phi^{b\beta}\pi^{a\alpha}_{b\beta}u^{a\alpha}$ $= \Pi^{b\beta}_{c\gamma}\Pi^{a\alpha}_{b\beta}\phi^{a\alpha}u^{a\alpha}$ proving the equality in question. We have shown that $[B^n_2(K^a \times K^\alpha \bmod C^{a\alpha}); \Pi^{a\alpha}_{b\beta}]$ is a spectrum. Let the direct limit of this spectrum be denoted by $B^n_2(R_1 \times R_2)$.

Let $\bar{u}^s \in B^s(R_1)$ and $\bar{u}^\sigma \in B^\sigma(R_2)$ have orders $\bar{e}^s \neq 0$ and $\bar{e}^\sigma \neq 0$. We shall associate with these two bundles an element $(\bar{u}^s, \bar{u}^\sigma)$ $\in B^{s+\sigma-1}_2(R_1 \times R_2)$. Consider a cohomology class of the coset $(u^{sa}, u^{\sigma\alpha})$, $u^{sa} \in \bar{u}^s$, $u^{\sigma\alpha} \in \bar{u}^\sigma$, u^{sa} of order $\bar{e}^s$, and $u^{\sigma\alpha}$ of order $\bar{e}^\sigma$. In view of (3) this cohomology class contains a cocycle which may be written

$$\frac{1}{(\bar{e}^s, \bar{e}^\sigma)}(f^{s-1a} \times f^{\sigma-1\alpha}). \tag{17}$$

with $f^{s-1a} = \bar{e}^s z^{sa}$, $z^{sa} \in u^{sa}$, and similar relations with superscript σ. We have

$$\sigma^{a\alpha}_{b\beta}(17) = \frac{1}{(\bar{e}^s, \bar{e}^\sigma)}(\sigma^a_b f^{s-1a} \times \sigma^\alpha_\beta f^{\sigma-1\alpha}). \tag{18}$$

because of (11) and the fact that the operation of forming coboundary commutes with $\sigma^{a\alpha}_{b\beta}$.[12] But $(\sigma^a_b f^{s-1a})^\cdot = \bar{e}^s \sigma^a_b z^{sa}$ with the same relation for superscript σ. Hence (18) is in a cohomology class of the coset $(\pi^a_b u^{sa}, \pi^\alpha_\beta u^{\sigma\alpha})$. Letting $H(z)$ still denote the cohomology class that contains the cocycle z, we have using these results and (16) that $\Pi^{a\alpha}_{b\beta}(u^{sa}, u^{\sigma\alpha}) = \Pi^{a\alpha}_{b\beta}\phi^{a\alpha}H(17) = \phi^{b\beta}\pi^{a\alpha}_{b\beta}H(17) = \phi^{b\beta}H(18) = (\pi^a_b u^{sa}, \pi^\alpha_\beta u^{\sigma\alpha})$. Hence we have a unique bundle $(\bar{u}^s, \bar{u}^\sigma) \in B^{s+\sigma-1}_2(R_1 \times R_2)$ defined by

$$(u^{sa}, u^{\sigma\alpha}) \in (\bar{u}^s, \bar{u}^\sigma). \tag{19}$$

From (4) we obtain

$$(\bar{e}^s, \bar{e}^\sigma)(\bar{u}^s, \bar{u}^\sigma) = 0. \tag{20}$$

Furthermore, let the sums $\sum p^s_i \bar{u}^s_i$ and $\sum p^\sigma_j \bar{u}^\sigma_j$ satisfy the conditions on the sums in (5). From (5) we have a relation which we label (21). The relation (21) is obtained from (5) by replacing each u by $\bar{u}$.

THEOREM 4. *The elements* $(\bar{u}^s, \bar{u}^\sigma)$, $s+\sigma-1=n$, *generate* $B^n_2(R_1 \times R_2)$, *and relations* (20) *and* (21) *are a complete set for these generators.*

Theorem 4 is a consequence of Theorem 2 and the definition of

[12] See [1, Theorem 2.221].

$B_2^n(R_1 \times R_2)$ as a direct limit. The proof is similar to the proof of Theorem 3.

THEOREM 5. *We have* $B_2^n(R_1 \times R_2) = B^n(R_1 \times R_2)/B_1^n(R_1 \times R_2)$.

PROOF. It follows from (16) that there is a homomorphic mapping of $B^n(R_1 \times R_2)$ upon $B_2^n(R_1 \times R_2)$.[12] The kernel of this homomorphism consists of the bundles of $B^n(R_1 \times R_2)$ which contain elements of the groups $B_1^n(K^\alpha \times K^\alpha \bmod C^{\alpha\alpha})$.[13] From (12) and (18) we see that this kernel is precisely $B_1^n(R_1 \times R_2)$.

REFERENCES

1. P. Alexandroff, *General combinatorial topology*, Trans. Amer. Math. Soc. vol. 49 (1941) pp. 41–105.

2. H. Freudenthal, *Eine Simplizialzerlegung des Cartesischen Productes Zweier Simplexe*, Fund. Math. vol. 29 (1937) pp. 138–144.

3. S. Lefschetz, *Algebraic topology*, Amer. Math. Soc. Colloquium Publications vol. 27 (1942) pp. 139–143.

PURDUE UNIVERSITY

[12] See [1, Theorems 3.41 and 3.411].

ON THE AVERAGE NUMBER OF REAL ROOTS OF A RANDOM ALGEBRAIC EQUATION

M. KAC

1. **Introduction.** Consider the algebraic equation

$$X_0 + X_1 x + X_2 x^2 + \cdots + X_{n-1} x^{n-1} = 0, \tag{1}$$

where the X's are independent random variables assuming real values only, and denote by $N_n = N(X_0, \cdots, X_{n-1})$ the number of real roots of (1). We want to determine the mean value (mathematical expectation = m.e.) of N_n when all X's have the same normal distribution with density

$$e^{-u^2}/\pi^{1/2}. \tag{2}$$

This problem was treated by Littlewood and Offord[1] who also considered the cases when the X's are uniformly distributed in $(-1, 1)$ or assume only the values $+1$ and -1 with equal probabilities. Littlewood and Offord obtain in each case the estimate

$$\text{m.e. } \{N_n\} \leqq 25(\lg n)^2 + 12 \lg n, \qquad n \geqq 2000.$$

In our case of normally distributed X's we shall be able to prove the exact formula

$$\text{m.e. } \{N_n\} = \frac{4}{\pi} \int_0^1 \frac{[1 - n^2[x^2(1 - x^2)/(1 - x^{2n})]^2]^{1/2}}{1 - x^2}\, dx \tag{3}$$

and then obtain the asymptotic relation

$$\text{m.e. } \{N_n\} \sim (2/\pi) \lg n \tag{4}$$

and the estimate

$$\text{m.e. } \{N_n\} \leqq (2/\pi) \lg n + 14/\pi, \qquad n \geqq 2. \tag{5}$$

In case the X's are not normally distributed (but all have the same distribution with standard deviation 1) one can still prove (4). The necessary limiting processes can then be carried out by using the central limit theorem of the calculus of probability and, as one may expect, the computations will be quite lengthy. On the other hand they will contribute relatively little to the general picture and, what

Presented to the Society, September 10, 1942; received by the editors July 2, 1942.

[1] J. London Math. Soc. vol. 13 (1938) pp. 288–295. No proofs are given in this paper, and the present author was unable to find them anywhere in print.

is worse, they may darken it by technicalities. We shall therefore assume in what follows that the X's are normally distributed with density given by formula (2).

2. **A formula for the number of real roots.** Let $f(x)$ be a continuous function in $(-\infty, \infty)$ having a continuous first derivative $f'(x)$ and only a finite number of turning points in each finite interval.

Let $\psi_\epsilon(x)$ be 1 if $-\epsilon<x<\epsilon$ and 0 otherwise. We then have the following lemma.

LEMMA 1. *If neither a nor b is a zero of $f(x)$, then for sufficiently small ϵ's*

$$(1/2\epsilon)\int_a^b \psi_\epsilon(f(x))\,|\,f'(x)\,|\,dx$$

is equal to the number of zeros of $f(x)$ inside the interval (a, b). (Multiple roots are counted only once.)[2]

We first notice that the set E_ϵ of those x's for which $\psi_\epsilon(f(x))=1$ is an open set and is therefore a sum of disjoint open intervals $I_1, I_2, I_3, \cdots$.

We choose ϵ small enough so that (a) no turning point of $f(x)$ in (a, b) lies in the strip $-\epsilon<y<\epsilon$ unless it happens to be at the same time a zero of $f(x)$; (b) no I_i includes either a or b.

Let $I_1, I_2, \cdots, I_r$, say, be totally contained in (a, b). Then, since it is easily seen that

$$\int_{I_i} |\,f'(x)\,|\,dx = 2\epsilon, \qquad 1 \leqq i \leqq r,$$

we have

$$(1/2\epsilon)\int_a^b \psi_\epsilon(f(x))\,|\,f'(x)\,|\,dx = (1/2\epsilon)\sum_{i=1}^{r}\int_{I_i} |\,f'(x)\,|\,dx = r.$$

This proves Lemma 1.

REMARK 1. An easy extension of the above reasoning gives for sufficiently small ϵ's

$$(1/2\epsilon)\int_a^b \psi_\epsilon(f(x))\,|\,f'(x)\,|\,dx = \text{number of zeros of } f(x) \text{ inside } (a, b) + \eta,$$

where $\eta = 0, 1/2, 1$ according as none, one or both of the numbers a, b are zeros of $f(x)$.

REMARK 2. If $f(x)$ is a polynomial then for sufficiently small ϵ's

[2] If $f(x)\equiv 0$ we consider $f(x)$ as not having any roots.

$$(1/2\epsilon)\int_{-\infty}^{+\infty}\psi_\epsilon(f(x))\,|\,f'(x)\,|\,dx$$

is equal to the number of real roots of $f(x)$. This is a trivial consequence of Lemma 1. One might notice that the limits of integration are really finite since $\psi_\epsilon(f(x))=0$ for sufficiently large $|\,x\,|$.

REMARK 3. The choice of ϵ in Remark 2 obviously depends on the coefficients of the polynomial. However, we can eliminate ϵ by restating the result of Remark 2 in the form:

$$\lim_{\epsilon\to 0}(1/2\epsilon)\int_{-\infty}^{+\infty}\psi_\epsilon(f(x))\,|\,f'(x)\,|\,dx = \text{number of real roots of } f(x).$$

If we put now

$$f(x) = X_0 + X_1x + \cdots + X_{n-1}x^{n-1},$$

we get

$$N_n = \lim_{\epsilon\to 0}(1/2\epsilon)\int_{-\infty}^{+\infty}\psi_\epsilon(f(x))\,|f'(x)\,|\,dx. \tag{6}$$

We shall need this lemma in what follows.

LEMMA 2. *If $f(x)$ is a polynomial of degree $n-1$, then for every $\epsilon>0$*

$$(1/2\epsilon)\int_{-\infty}^{+\infty}\psi(f(x))|\,f'(x)|\,dx \leqq 3n-5.$$

In this case the set E (see proof of Lemma 1) is a sum of at most $2n-3$ open intervals. Indeed, each I_i contains either a real root of $f(x)$ or a turning point and there are at most $n-1$ real roots and at most $n-2$ turning points.

The proof of Lemma 2 follows now from the obvious remark that

$$\int_{I_i}|\,f'(x)\,|\,dx \leqq 2\epsilon(m_i+1), \qquad 1\leqq i\leqq 2n-3,$$

where m_i is the number of turning points in I_i.

3. **Interchange of two limiting processes.** We now reduce the computation of m.e. $\{N_n\}$ to the computation of m.e. $\{\psi_\epsilon(f(x))|f'(x)|\}$ by means of this lemma.

LEMMA 3.

$$\text{m.e. }\{N_n\} = \lim_{\epsilon\to 0}(1/2\epsilon)\int_{-\infty}^{+\infty}\text{m.e. }\{\psi_\epsilon(f(x))\,|\,f'(x)\,|\}dx.$$

Random variables can be considered as measurable functions defined on a set Ω with a completely additive Lebesgue measure. (The measure of Ω is 1.) Mathematical expectation (m.e.) is nothing but a Lebesgue integral with respect to that measure. Both N_n and $\psi_\epsilon(f(x))|f'(x)|$ are then measurable functions on Ω and they can be represented symbolically as

$$N_n(\mu) \quad \text{and} \quad g_\epsilon(x, \mu),$$

x and ϵ being real parameters. We also have

$$\text{m.e. } \{N_n\} = \int_\Omega N_n(\mu)d\mu,$$

$$\text{m.e. } \{g_\epsilon(x, \mu)\} = \int_\Omega g_\epsilon(x, \mu)d\mu,$$

$d\mu$ indicating that integration is being performed with respect to the Lebesgue measure in Ω. We first notice that

$$\text{(7)} \qquad \text{m.e.} \left\{ \int_{-\infty}^{+\infty} g_\epsilon(x, \mu)dx \right\} = \int_{-\infty}^{+\infty} \text{m.e. } \{g_\epsilon(x, \mu)\}dx.$$

This follows from Fubini's theorem if one notices that the integral with respect to dx is really an integral between finite limits (depending, of course, on ϵ and μ). Furthermore, by Lemma 2 for every $\epsilon > 0$

$$(1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu)dx < 3n - 5,$$

and hence, by a well known theorem from Lebesgue's theory,

$$\lim_{\epsilon \to 0} \int_\Omega \left[(1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu)dx \right] d\mu$$

$$= \int_\Omega \left[\lim_{\epsilon \to 0} (1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu)dx \right] d\mu.$$

This when combined with (6) and (7) completes the proof of Lemma 3.

4. **A formula for** m.e. $\{N_n\}$. In order to compute m.e. $\{\psi_\epsilon(f(x))|f'(x)|\}$ we need the following lemma.

LEMMA 4. *If* $\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, \beta_0, \beta_1, \cdots, \beta_{n-1}$ *are real numbers,* $\sum \alpha_i^2 = \alpha, \sum \beta_i^2 = \beta, \sum \alpha_i \beta_i = \gamma$ *and if* $\Delta = \alpha\beta - \gamma^2 > 0$, *then the density of the joint distribution of* $\alpha_0 X_0 + \cdots + \alpha_{n-1} X_{n-1}$ *and* $\beta_0 X_0 + \cdots + \beta_{n-1} X_{n-1}$ *is equal to*

$$(1/\pi\Delta^{1/2}) \exp \{-(\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta\}.$$

This fact is well known.[3]

It also follows from well known facts that

$$\text{m.e.} \{\psi_\epsilon(\alpha_0 X_0 + \cdots + \alpha_{n-1}X_{n-1}) \mid \beta_0 X_0 + \cdots + \beta_{n-1}X_{n-1}\}$$
$$= (1/\pi\Delta^{1/2}) \int_{-\infty}^{+\infty} \psi_\epsilon(u) \mid v \mid \exp \{-(\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta)\} du dv.$$

Let

$$\int_{-\infty}^{+\infty} \mid v \mid \exp \{-(\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta)\} dv = F(u),$$

then $F(u)$ is a continuous function of u and we get

$$\lim_{\epsilon \to 0} (1/2\epsilon) \text{ m.e.} \{\psi_\epsilon(\sum \alpha_i X_i) \mid \sum \beta_i X_i \mid\}$$
$$= \lim_{\epsilon \to 0} (1/2\epsilon\pi\Delta^{1/2}) \int_{-\epsilon}^{\epsilon} F(u) du = F(0)/\pi\Delta^{1/2}.$$

But

$$F(0) = \int_{-\infty}^{+\infty} \mid v \mid \exp\{-\alpha v^2/\Delta\} dv = \Delta/\alpha$$

and finally

$$\lim_{\epsilon \to 0} (1/2\epsilon) \text{ m.e.} \{\psi_\epsilon(\sum \alpha_i X_i) \mid \sum \beta_i X_i \mid\} = \Delta^{1/2}/\alpha.$$

If we now put $\alpha_0 = 1$, $\alpha_1 = x$, $\cdots$, $\alpha_{n-1} = x^{n-1}$, $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2x$, $\cdots$, $\beta_{n-1} = (n-1)x^{n-2}$, we obtain by an elementary computation

$$\Delta = \frac{x^{4n} - n^2 x^{2(n+1)} + 2(n^2 - 1)x^{2n} - n^2 x^{2(n-1)} + 1}{(x^2 - 1)^4}.$$

It is easy to prove that $\Delta > 0$ for every real x and by combining the considerations of this section with Lemma 3 we obtain

$$\text{m.e.} \{N_n\}$$
$$(8) \quad = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{(x^{4n} - n^2 x^{2(n+1)} + 2(n^2 - 1)x^{2n} - n^2 x^{2(n-1)} + 1)^{1/2}}{(x^2 - 1)^2(1 + x^2 + x^4 + \cdots + x^{2n-2})} dx.$$

[3] See for instance S. Bernstein, *Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes*, Math. Ann. vol. 97 (1927) pp. 2–59, in particular, chap. 3 (pp. 43–59). This chapter contains limit theorems by means of which one can handle the case of not normally distributed X's. Lemma 4 is but a simple consequence of the main results of that chapter.

5. **Estimates and the asymptotic formula.** By an elementary transformation we get

$$\text{m.e. } \{N_n\} = \frac{4}{\pi}\int_0^1 \frac{(1-h_n^2(x))^{1/2}}{1-x^2}\,dx,$$

where

$$h_n(x) = \frac{nx^{n-1}(1-x^2)}{1-x^{2n}}.$$

It should be mentioned that the value of the integrand at $x=1$ is $(n^2-1/12)^{1/2}$. We have

$$1-x^{2n} = (1-x)(1+x+x^2+\cdots+x^{2n-1}) < 2n(1-x)$$

and therefore

$$\begin{aligned} h_n(x) &> x^{n-1}(1+x)/2 \\ 1-h_n^2(x) &< [1-x^{n-1}(1+x)/2][1+x^{n-1}(1+x)/2] \\ &< 2-x^{n-1}(1+x). \end{aligned}$$

Using the mean value theorem we get

$$\begin{aligned} 2-x^{n-1}(1+x) &= (1-x)[\theta^{n-1}+(n-1)\theta^{n-2}(1+\theta)] \\ &< (2n-1)(1-x), \qquad x<\theta<1. \end{aligned}$$

Finally,

$$(1-h_n^2(x))^{1/2}/(1-x^2) < (2n-1)^{1/2}/(1-x)^{1/2}, \qquad 0 \leqq x < 1.$$

On the other hand

$$(1-h_n^2(x))^{1/2}/(1-x^2) \leqq 1/1-x^2, \qquad 0 \leqq x < 1,$$

and we can write

$$\int_0^1 \frac{(1-h_n^2(x))^{1/2}}{1-x^2}\,dx < \int_0^{1-1/n} \frac{dx}{1-x^2} + \int_{1-1/n}^1 \frac{(2n-1)^{1/2}}{(1-x)^{1/2}}\,dx$$

$$= (1/2)\lg n + (1/2)\lg(2-1/n) + 2(2-1/n)^{1/2} < (1/2)\lg n + 3\cdot 5.$$

Finally,

$$\text{m.e. } \{N_n\} < (2/\pi)\lg n + 14/\pi, \qquad n \geqq 2. \tag{9}$$

In order to obtain a lower estimate let ϵ and δ be arbitrary positive numbers less than 1. We have

$$\int_0^1 \frac{(1-h_n^2(x))^{1/2}}{1-x^2}\,dx > \int_0^{1-n^{\delta-1}} \frac{(1-h_n^2(x))^{1/2}}{1-x^2}\,dx.$$

But for $0 \leqq x \leqq 1-n^{\delta-1}$

$$h_n(x) < nx^{n-1} \leqq n(1-n^{\delta-1})^{n-1}$$

and this last expression can be made smaller than $\epsilon^{1/2}$ if n is sufficiently large. Hence, for sufficiently large n,

$$\int_0^1 \frac{(1-h_n^2(x))^{1/2}}{1-x^2}\,dx > \int_0^{1-n^{\delta-1}} \frac{(1-\epsilon)^{1/2}}{1-x^2}\,dx$$

$$> \frac{(1-\epsilon)^{1/2}(1-\delta)}{2} \lg n.$$

Using (9) and the fact that ϵ and δ can be made arbitrarily small we obtain the asymptotic formula

$$\text{m.e.}\,\{N_n\} \sim (2/\pi) \lg n. \tag{10}$$

Let us add that it follows from (9) that the probability that (1) has more than $\sigma \lg n$ real roots is less than

$$2/\pi\sigma + 14/\pi\sigma \lg n.$$

This result is not trivial only in case $\sigma > 2/\pi$.

6. **Final remarks.** It is quite clear that the average number of real roots of (1) falling in the interval (a, b) is given by formula (8) if one replaces $-\infty$ and $+\infty$ by a and b, respectively. The proof of this statement does not differ from the one given above and one must only notice that the probability that either a or b is a root of (1) is 0.

One can also see almost immediately that if (a, b) does not contain either 1 or -1 the average number of real roots of (1) falling within (a, b) is $O(1)$. This means, roughly speaking, that most of the real roots of most of the equations cluster around 1 and -1. The problem of the exact determination of the average distribution of real roots of (1) on the real axis will, of course, depend on a more delicate treatment of the integral (8). It would also be interesting to know the higher moments of N_n. The present line of attack would lead to very complicated integrals, but it may be hoped that some other approach will furnish more information about the distribution of the number of real roots of equation (1).

CORNELL UNIVERSITY